SEISMIC THRESHOLD DETERMINATION

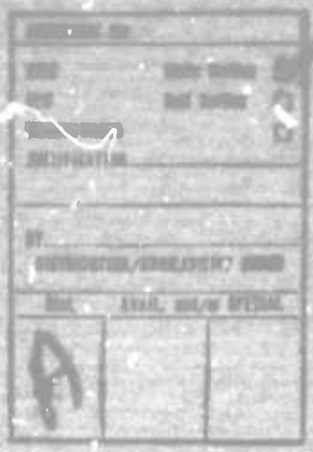
David von Seggern, et al

Teledyne Geotech

Prepared for:

Defense Advanced Research Projects Agency
Air Force Technical Applications Center

22 March 1974



Disclaimer: Neither the Defense Advanced Research Projects Agency nor the Air Force Technical Applications Center will be responsible for information contained herein which has been supplied by other organizations or contractors, and this document is subject to later revision as may be necessary. The views and conclusions presented are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, the Air Force Technical Applications Center, or the US Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1 REPORT NUMBER SDAC-TR-74-3	2 GOVT ACCESSION NO.	3 RECIPIENT'S CATALOG NUMBER AD-A010 143
4 TITLE (and Subtitle) SEISMIC THRESHOLD DETERMINATION		5 TYPE OF REPORT & PERIOD COVERED Technical 6 PERFORMING ORG. REPORT NUMBER
7 AUTHOR(s) von Seggern, David; Blandford, Robert		8 CONTRACT OR GRANT NUMBER(s) F08606-74-C-0006
9 PERFORMING ORGANIZATION NAME AND ADDRESS Teledyne Geotech 314 Montgomery Street Alexandria, Virginia 22314		10 PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11 CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency Nuclear Monitoring Research Office 1400 Wilson Blvd.-Arlington, Va. 22209		12 REPORT DATE 22 March 1974 13 NUMBER OF PAGES 179
14 MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) VELA Seismological Center 312 Montgomery Street Alexandria, Virginia 22314		15 SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16 DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17 DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18 SUPPLEMENTARY NOTES

19 KEY WORDS (Continue on reverse side if necessary and identify by block number)

Seismic threshold
Monte Carlo simulation
M_s-m_b relation
Correlated signals
Correlated noise
Detection probabilities

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

Problems in accurate determination of seismic threshold magnitude are analytically tractable for only some simple, but interesting, cases. This report gives these analytic results and proceeds to a simulation experiment to provide insights where analytic predictions are not possible. Differences in direct, incremental, and cumulative threshold determinations; distinctions between true

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SEISMIC THRESHOLD DETERMINATION

SEISMIC DATA ANALYSIS CENTER REPORT NO.: SDAC-TR-74-3

AFTAC Project No.:	VELA VT/4709
Project Title:	Seismic Data Analysis Center
ARPA Order No.:	1620
ARPA Program Code No.:	3F10
Name of Contractor:	TELEDYNE GEOTECH
Contract No.:	F08606-74-C-0006
Date of Contract:	01 July 1973
Amount of Contract:	$2,152,172
Contract Expiration Date:	30 June 1974
Project Manager:	Royal A. Hartenberger (703) 836-3882

P. O. Box 334, Alexandria, Virginia 22313

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

ABSTRACT

Problems in accurate determination of seismic threshold magnitude are analytically tractable for only some simple, but interesting, cases. This report gives these analytic results and proceeds to a simulation experiment to provide insights where analytic predictions are not possible. Differences in direct, incremental, and cumulative threshold determinations; distinctions between true and observed thresholds; effects of correlated signals and noise; feedback between LR and P thresholds; and estimation of M_s-m_b relationships are treated in this simulation experiment. A comparison of simulated detection results is made with real data from LASA on the Kamchatka-Kurile region and good agreement is obtained between predictions and observations.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

INTRODUCTION

Determination of seismic threshold for a station or network of stations serves to evaluate that station or network in the context of seismic detection. Threshold level reflects several different aspects of the station or network: seismic background noise, relative signal strength, seismometry, recording equipment, processing techniques, and analysis procedures. If one aspires to quantify these effects or to evaluate station and network detection capability, it is then necessary to make an accurate determination of threshold. It is also desirable to have a technique for predicting thresholds and for comparing predictions to observations.

Determining seismic thresholds has several facets in practice. An empirical threshold can be determined directly by plotting the percentage of events detected in each equal magnitude increment against the known magnitudes. This is the direct method; it gives an unbiased threshold only if one has an independent source of magnitude information whose threshold is much lower than the station or network under consideration and whose reported magnitudes have a small variance and are insignificantly biased. Since an independent source of such nature is often unavailable, one must use a method dependent upon magnitude information from the network or station under consideration; this involves plotting the number of events having a certain phase

(e.g., LR) detected against the magnitudes calculated from that phase (e.g., M_S). This is the incremental method; thresholds are found by estimating the true seismicity-magnitude relation from the asymptotic part of the data at high magnitudes or by a maximum-likelihood method (Kelley and Lacoss, 1969). We will show that the use of observed, rather than true or independent, magnitudes distorts the simple picture presented by the direct method and effects the threshold determination. A variation of the incremental method is to plot the cumulative number of detections (with magnitude greater than or equal to a given observed magnitude) against the observed magnitudes; this is the cumulative method. We will show that this approach, taken to smooth the data points, heightens the threshold distortion even more.

Another somewhat different approach is used in seismology and is a type of direct threshold determination. Here, in the manner of the direct method, the percentage of detections of surface waves in a given body-wave magnitude increment is plotted against the body-wave magnitude, supplied by a source which presumably detects body waves much more frequently than the station or network under consideration detects surface waves. Thresholds are then converted from the m_b to M_S scale. We will show how this method of M_S threshold determination is so fraught with pitfalls that only the most careful attention to detail can produce reliable thresholds.

The matter of calculated thresholds, by use of program NETWORTH (Wirth, 1970), also deserves attention since the authors have noted cases of unacceptable discrepancy between calculated and observed thresholds. All the discrepancy is not necessarily due to faulty empirical threshold determination, and we must reconsider the assumptions in the common scheme for calculating station or network thresholds.

This report discusses several of the important aspects of threshold determination, calculated and observed, which we feel have not been given exposure or properly interrelated; we feel that attention to these aspects, along with those presented by previous investigators, will result in unbiased threshold determinations which are suitable for use in comparisons between stations and between networks.

We must make some simplifying assumptions about certain phenomena in order to make the mathematics tractable; they are, however, realistic and commonly used. Therefore, we allow the signal amplitudes from an event to be lognormally distributed (Freedman, 1967; von Seggern, 1972). The same applies to the noise amplitudes at a station (Geotech, 1966). We allow the logarithm of the number of events which occur in a specified time interval to be proportional to magnitude (Gutenberg and Richter, 1954; Scholz, 1968); that is, we use a linear seismicity-amplitude relation throughout. We also assume a linear relationship between M_s and m_b everywhere.

The first parts of this report establish the concepts, and theoretical derivations are given when necessary. A later part of this report describes the method and results of a simulation experiment which serves to verify the predictions of the first parts and to give insights where mathematical predictions are not attainable. The techniques used in this simulation are essential tools for satisfactory evaluations of practical networks. The last part of the report deals with refinements in calculating seismic thresholds.

To aid the reader, we provide next for reference a list of notations appearing in equations throughout the report and also a glossary of important recurrent terms used in the report.

NOTATION

μ	-	mean common logarithm of the noise amplitude at a station
σ_n	-	standard deviation of the common logarithm of noise amplitude at a station
A	-	calculated amplitude of the signal at a given station
$\hat{A}$	-	observed amplitude of the signal at a given station
$\tilde{m}_b$	-	true body-wave magnitude
$\tilde{M}_s$	-	true surface-wave magnitude
m_b	-	operational body-wave magnitude
M_s	-	operational surface-wave magnitude
$\hat{m}_b$	-	observed body-wave magnitude
$\hat{M}_s$	-	observed surface-wave magnitude
σ_m	-	standard deviation of operational magnitudes about the true magnitude
σ_{mb}	-	standard deviation of operational body-wave magnitudes about true magnitudes
σ_{ms}	-	standard deviation of operational surface-wave magnitudes about true magnitudes
σ_s	-	standard deviation of observed magnitudes about operational magnitude
σ_{sb}	-	standard deviation of observed body-wave magnitudes about operational magnitudes
σ_{ss}	-	standard deviation of observed surface-wave magnitudes about operational magnitudes

$N(m)$	-	number of events occurring with magnitude m
$N_d(m)$	-	number of events detected with magnitude m
α	-	slope of the base 10 seismicity-magnitude relation
α'	-	slope of the base e seismicity-magnitude relation
β	-	intercept of the seismicity-magnitude relation
β'	-	10^{β}
a	-	slope of the M_s-m_b relation
b	-	intercept of the M_s-m_b relation
Z	-	normal probability density function
Φ	-	unit normal probability distribution function
erf	-	error function
$P_D(m)$	-	direct probability of detection curve
$P_I(m)$	-	incremental probability of detection curve
$P_C(m)$	-	cumulative probability of detection curve
$P(x)$	-	probability distribution function
$P(x\|y)$	-	conditional probability distribution function
z_α	-	percentage point for the unit normal distribution
$B_{k,n}$	-	binomial probability for exactly k of n occurrances
$f(x)$	-	probability density function
$f(x\|y)$	-	conditional probability density function
ρ	-	correlation coefficient

GLOSSARY

Bias

source bias - a difference between the operational magnitude and the true magnitude due to differences in fault dimensions and shape, existing stress levels, close-in elastic parameters, etc., which are not explicitly allowed for in the network's routine determinations of magnitude.

path-station bias - a difference between the observed magnitude and the operational magnitude due to differences in depth, radiation pattern, path effects, station characteristics, etc., whose effects may be removed from the magnitude estimates by the network's routine analysis.

threshold bias - a difference between observed and true threshold magnitudes resulting from incremental and cumulative threshold determination methods.

Magnitude

true magnitude - the scale which most nearly represents the energy release of a seismic event; there is a one-to-one mapping between this scale and e.g., "yield" for explosions or "seismic moment" for earthquakes.

operational magnitude - tied to a particular phase such as $P(m_b)$ or $LR(M_s)$, this is the magnitude which an ideal network (in which a large number of stations with extremely low threshold densely cover the globe), using its routine highly-refined magnitude estimation techniques, would estimate for a seismic event.

observed magnitude - the magnitude of a seismic event as routinely determined by a station or real network (in which a limited number of stations with varying thresholds sparsely cover the globe).

Correlation

correlated noise - noise with non-zero correlation of changes in long-term ($\sim$ 1 hour) levels among stations of a network.

correlated signals - correlated differences among stations between the observed station magnitudes and true magnitude. These differences are related to source bias.

Relation

seismicity-magnitude relation - the equation expressing the dependence of the number of events occurring within a specified time on magnitude, usually of the form log(number) = α(magnitude) + β.

M_s-m_b relation - the equation expressing the dependence of surface-wave magnitude on body wave magnitude or vice-versa, usually of the form $M_s = am_b + b$.

Threshold

threshold magnitude - a suitably chosen magnitude at which a certain proportion, usually 50% or 90%,

of the events of interest are detected via a phase such as P or LR.

direct threshold determination - a method of threshold magnitude determination whereby the percentage of events for which a phase of interest (e.g., P or LR) was detected is plotted against some magnitude which is not calculated from those phases themselves.

incremental threshold determination - a method of threshold magnitude determination whereby the number of event detections via a certain phase (e.g., P or LR) in each magnitude increment is plotted against magnitudes determined usually, but not always, from these phases themselves. Thresholds are determined where plotted points begin to deviate from the straight-line approximation to the seismicity-magnitude relation, which is determined by higher-magnitude events where 100% detection is achieved.

cumulative threshold determination - same as incremental threshold determination, except that cumulative numbers of detections with magnitudes greater than or equal to the current magnitude are plotted.

CONCEPT OF SEISMIC MAGNITUDE

In discussing seismic thresholds, it is necessary to view seismic magnitude quite precisely. We have a need for three conceptions of magnitude: true, operational and observed, as defined below. Such a trichotomy has been implied for explosion magnitudes by von Seggern (1973). We must then discuss our theoretical predictions and practical applications in this framework and clearly distinguish at all times which magnitude concept we are using.

The question of the "true" magnitude of a seismic event is a difficult one for which no completely satisfying answer is available. One approach to this problem is to regard the magnitude which maps one-to-one into the radiometric or design yield of an explosion as the fundamental, or as we shall say the "true", magnitude; for earthquakes, "true" magnitude could be linked to seismic moment or the total radiated energy.

Even for a perfect unbiased network, the "operational" magnitudes, such as m_b, M_S, or M_L, as computed by standard seismological procedures, would vary for events of the same true magnitude. This variation we call source bias. For example, difference in depths of burial or in properties of the surrounding media will result in different magnitudes for the same yield. Similarly, different fault mechanisms, depths of focus, and states of stress can result in varying operational magnitudes for otherwise identical earthquakes. Finally,

we define our real estimates of magnitude as the "observed" magnitudes; these will vary about the operational magnitude due to radiation patterns, station noise levels, path and station effects, etc. The true magnitudes are seismologically unmeasurable and therefore unavailable to use. Likewise operational magnitudes, although measurable in principle, are unavailable to use because of the limitations of real seismic networks. In general, the larger the event, the more closely careful analysis will cause the observed magnitude to approach the operational magnitude.

For presenting results it will be desirable to express thresholds in terms of one or more of the above three magnitudes, depending on the frame of reference of the discussions. For example, if one asks for the threshold of detection of a 2 kt shot in granite, the answer is different, as we shall show, from that of an event with operational m_b equal to that of the expected value of the operational m_b for 2 kt in granite. Of course for most cases of interest these differences are not great (on the order of 0.1 magnitude unit).

INCREMENTAL AND DIRECT DETECTION PROBABILITIES

Taking the assumptions on signal and noise amplitudes stated in the introduction, and further assuming that the signal only need exceed the noise for a detection to occur, the probability of detection of an event phase at a station is given by a normal distribution function:

$$P_D(m) = \Phi \left[\frac{\log A - \mu}{(\sigma_s^2 + \sigma_n^2)^{1/2}} \right] \qquad (1)$$

where μ is the mean logarithm of noise amplitude at the station, σ_s and σ_n are the standard deviations of the logarithms of signal and noise amplitudes respectively, and A is the calculated amplitude of that phase at the station from an event with operational magnitude m. Computation of $P_D(m)$ can be extended to a network if desired (Wirth, 1970). Blandford and Wirth (1973) have analyzed the case of a station or a network with automatic power or F detectors and have justified the use of (1) for those stations. The detection curve for a given station-region pair should take the form of equation (1) when the method of direct threshold determination is employed. Now, when a station is used to observe events for a period of time, data on numbers of events versus observed magnitude $\hat{m}$ at that station may be accumulated; and incremental thresholds can be determined from this data by maximum likelihood or least-squares (Kelley and Lacoss, 1969).

However, these thresholds are biased due to the fact that they are based on observed magnitudes, not operational magnitudes. We generally need to know $P_D(m)$, whereas the incremental method determines:

$$P_I(\hat{m}) = \Phi\left[\frac{\log\hat{A} - \mu}{\sigma_n}\right] \qquad (2)$$

where $\hat{A}$ is now the estimated amplitude associated with $\hat{m}$ which is m plus a random error variable normally distributed with zero mean and variance σ_s^2 due to path-station effects. The bias of the incremental threshold magnitude relative to the direct one can be explicitly calculated as follows. First consider the number of events detected at both m and $\hat{m}$:

$$N_d(m) = N(m) \cdot \Phi\left[\frac{m - \mu'}{(\sigma_s^2 + \sigma_n^2)^{1/2}}\right]$$

$$N_d(\hat{m}) = N(\hat{m}) \cdot \Phi\left[\frac{\hat{m} - \mu'}{\sigma_n}\right] .$$

A constant equal to the distance correction factor in magnitude calculation for a particular station-region pair has been absorbed into μ to get μ'. Taking the fraction of events detected in each case to be α percent and equating the detection probabilities gives:

$$\Phi\left[\frac{m_\alpha - \mu'}{(\sigma_s^2 + \sigma_n^2)^{1/2}}\right] = \Phi\left[\frac{\hat{m}_\alpha - \mu'}{\sigma_n}\right]$$

where m_α and $\hat{m}_\alpha$ are the operational and observed magnitudes at which α percent detection occurs. Equating the arguments to z_α in the above Φ functions, which are equivalently the unit normal density function integrated to Z_α, and eliminating μ' results in:

$$\hat{m}_\alpha = m_\alpha - z_\alpha \left[(\sigma_s^2 + \sigma_n^2)^{1/2} - \sigma_n \right]. \qquad (3)$$

This relation gives the biased observed threshold magnitude $\hat{m}_\alpha$ which is determined by the incremental method when the operational threshold lies at m_α. For example, if the operational threshold is chosen to be at 90% and $\sigma_s = \sigma_n = .3$, then $\hat{m}_{.9} \approx m_{.9} - 0.16$. The bias is symmetric about the 50% threshold; thus for the 10% operational threshold, $\hat{m}_{.1} \approx m_{.1} + 0.16$. Table I lists the 10%, 50%, and 90% threshold magnitude bias for various values of σ_s and σ_n. Note that the above discussion can be extended to true and observed magnitudes if σ_s^2 is replaced by $\sigma_s^2 + \sigma_m^2$, where σ_m^2 is the variance of the source effects about true magnitude. Since use of observed magnitudes gives a biased estimate of the operational or true 90% threshold but an unchanged estimate of the 50% threshold, in practice better comparisons of incremental thresholds among stations might well be made at the 50% level.

The threshold magnitude bias produced by a network of stations cannot be expressed in a simple equation such as that for a single station (equation 3).

However, if we can estimate numerically what the incremental detection curve on observed magnitude for a network should be, then by comparison with the incremental curve on operational magnitude, we can determine network threshold magnitude bias. Let us begin to estimate the incremental network detection curve with an expression from Herrin and Tucker (1972) for the probability density of observed $\hat{m}$ at a station for the subset of events with operational magnitude m detected at that station:

$$f(\hat{m}) = \left[\Phi\left(\frac{\hat{m} - \mu'}{\sigma_n}\right) / \Phi\left(\frac{m - \mu'}{\sigma}\right)\right] \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{1}{2}\left(\frac{\hat{m} - m}{\sigma_s}\right)^2\right] \quad (4)$$

where σ replaces $(\sigma_s^2 + \sigma_n^2)^{1/2}$. This differs from the probability density of all $\hat{m}$ at a station for all the operational magnitude m events (neglecting the fact that $\hat{m}$ cannot be calculated for undetected events) which must simply be normal with mean m and variance σ_s^2. We need the probability density of $\hat{m}$ averaged over those stations which detected each of the magnitude m events. This is a sum of probability densities similar to (4) for all the cases of "k of n" detection, each suitably weighted by the probability that exactly k of n stations detect.

Unfortunately, we were able to derive the distribution $f(\hat{m})$ only for the case of two stations of equal detection capability; this derivation is given in Appendix 1, and the resulting expression (A12) is:

$$f(\bar{\hat{m}}) = \frac{\exp\left[-\frac{1}{2}\left(\frac{\bar{\hat{m}} - m}{\sigma_s/2}\right)^2\right]}{\pi^{1/2}\sigma_s\left[\Phi\left(\frac{m - \mu'}{\sigma}\right)\right]^2}\left\{\left[\Phi\left(\frac{\bar{\hat{m}} - m}{\sigma'}\right)\right]^2 - \sum_{n=0}^{\infty}\frac{(-1)^n\left[Z(n)\left(\frac{\bar{\hat{m}} - m}{\sigma'}\right)\right]^2\left[\frac{\sigma_n^2}{\sigma'^2}\right]^{n+1}}{(n+1)!}\right\} \quad (5)$$

Although we will not demonstrate it here, the expected value of $\bar{\hat{m}}$ with the probability density given by (5) is the same as the expected value of $\hat{m}$ with the probability density given by (4); in other words, on the average, the expected observed magnitude bias, $\hat{m}$-m, for an event of given m will be the same regardless of whether one or several stations (of equal capability) detected it. (Of course the bias does vary as a function of m.) Herrin and Tucker (1972) have demonstrated this fact without having to derive the distribution of $f(\bar{\hat{m}})$.

To calculate the number of events with magnitude $\hat{m}$ that will appear for a two-station network or a single station, we integrate $f(\bar{\hat{m}})$ or $f(\hat{m})$, respectively, weighted by the number of events detected at m, over operational magnitude m:

$$N_d(\hat{m}) = \int_{-\infty}^{\infty} f(\hat{m}) \cdot P_D(m) \cdot N(m) \cdot dm \quad \text{(single-station)} \quad (6a)$$

$$N_d(\hat{m}) = \int_{-\infty}^{\infty} [f(\hat{m}) \cdot B_{1,2}(m) + f(\hat{m}) \cdot B_{2,2}(m)] \cdot N(m) \cdot dm \qquad (6b)$$

(two-station network)

where $B_{1,2}$ and $B_{2,2}$ are the binomial probabilities of one and two station detection, respectively, given by:

$$B_{k,n} = \binom{n}{k} p^n q^{n-k}$$

where p is given by equation (1) and q is 1-p. We will need for our derivation a relation between the number of events which occur and the magnitude; this is commonly expressed as:

$$\log_{10} N(m) = -\alpha m + \beta$$

or as

$$N(m) = \beta' e^{-\alpha' m} \qquad (7a)$$

where $\beta' = 10^{\beta}$ and $\alpha' = 2.303\alpha$. We performed the integration of (6a) and (6b) numerically for the case of $\sigma_n = .2$, $\sigma_s = .3$, $\mu = 3.3$, $\alpha = 0.8$, $\beta' = 1.7 \times 10^5$, over a finite range of m to get both a network incremental curve and a single-station incremental curve. The results are shown in Figure 1 along with the assumed seismicity-magnitude relation for operational magnitude given by $\alpha = 0.8$, $\beta' = 1.7 \times 10^5$. In this figure the seismicity-magnitude relation is plotted as a function of operational magnitude m, while the incremental curves are

plotted against observed magnitude $\hat{m}$. An incremental threshold determination for the single-station case is made by observing the percentage falloff in detection of the number of possible events predicted by fitting a straight line to the single-station points at higher magnitudes. The single-station 50% and 90% threshold magnitude on operational magnitude, as calculated by equation (1), are indicated for reference. Note that the single-station bias for the 90% threshold magnitude is toward a lesser magnitude as predicted by equation (3) and that the bias for the 50% threshold magnitude is zero as also predicted by equation (3). The hump in the network incremental curve near the threshold is characteristic and is confirmed by points generated in a simulation of the two-station network by a computer program which we will discuss later. This hump is an unfortunate happenstance since it will degrade our capability to make threshold determinations for networks by fitting a straight line to the observed seismicity-magnitude data, $N_d(\hat{m})$. Consequently, we cannot easily determine the threshold bias for a network because of the uncertainty about the true asymptotic slope of the seismicity-magnitude relation when given a display of network data such as that in Figure 1.

When the number of detected events $N_d(\hat{m})$ is plotted against observed magnitude, then the asymptotic portion of this curve where $N_d(\hat{m}) \approx N(\hat{m})$ is displaced from the seismicity-magnitude relation plotted against true or operational magnitude as in Figure 1. This displacement

is due to the fact that more events from a lower operational magnitude scatter into a given observed magnitude than from a higher operational magnitude. We assume that the observed magnitudes are equal to an operational magnitude plus a random variable normally distributed with zero mean and variance σ_s^2. We can calculate the displacement of the asymptotic portion of the observed seismicity-magnitude line relative to the real one by finding the number of events with observed $\hat{m}$:

$$N(\hat{m}) = \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}\sigma_s} \exp\left[-\frac{1}{2}\left(\frac{\hat{m} - m}{\sigma_s}\right)^2\right] \cdot N(m) \cdot dm.$$

Substituting (7a) in the above gives

$$N(\hat{m}) = \frac{\beta'}{(2\pi)^{1/2}\sigma_s} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{\hat{m} - m}{\sigma_s}\right)^2\right] e^{-\alpha' m} dm$$

which is easily integrated to give:

$$N(\hat{m}) = \beta' e^{-\alpha' \hat{m}} e^{\alpha'^2 \sigma_s^2/2}. \qquad (7b)$$

Let $\hat{m}_k$ and m_k be the observed and operational magnitudes at which k events appear; then setting $N(\hat{m}_k)$ equal to $N(m_k)$ and taking the natural logarithm, we have:

$$\hat{m}_k = m_k + \frac{\alpha' \sigma_s^2}{2} \qquad (8)$$

The horizontal displacement of the asymptotic portion of the seismicity-magnitude line is just $\alpha' \sigma_s^2/2$, which

should be roughly .1 to .2 magnitude units for a single station for reasonable values of α and σ_s^2. We see in Figure 1 that the single-station observed seismicity line has been displaced relative to the real one .10 outward along the magnitude scale. The displacement of the line vertically leads to an overestimation of seismicity at a given magnitude, given by the factor $10^{\alpha\alpha'\sigma_s^2/2}$, which is approximately 1.15 for the single-station case in Figure 1. For a network, σ_s^2 is divided by the number of stations over which magnitude is averaged as reflected in the lesser displacement of the two-station line in Figure 1; and the displacement of the observed network seismicity-magnitude line from the true one should be insignificant for a network of more than a few stations.

The above derivation on displacement of the seismicity line can likewise be made for operational versus true magnitudes simply by replacing σ_s by σ_m. However, if there is no source bias, i.e. $\sigma_m=0$, then the seismicity line based on operational magnitude will not differ from that based on true magnitude, and the observed network seismicity line will approach that based on true magnitude.

In this section we have shown that thresholds (other than 50%) determined incrementally for a single station will differ from thresholds based on operational or true magnitude, that a hump will appear in the incremental curve for a network, thus making threshold determination in the network case difficult,

and that seismicity can be somewhat overestimated by incremental detection plots. We will next investigate some of the problems which arise in connection with the use of cumulative plots to determine threshold.

CUMULATIVE AND INCREMENTAL DETECTION PROBABILITIES

Often it seems necessary to determine thresholds by plotting the cumulative numbers of events with magnitudes greater than a given magnitude instead of the incremental numbers as in the previous section. Clearly, this method utilizes the same data as the incremental method only in a modified fashion. As Lacoss (1972) has already pointed out, cumulatively determined threshold magnitudes will always be lower than incrementally determined ones. We can derive this bias analytically as follows. First, we express the incremental and cumulative curves as:

$$P_I(\hat{m}) = N_d(\hat{m})/N(\hat{m}) \tag{9}$$

$$P_C(\hat{m}) = \int_{\hat{m}}^{\infty} N_d(\hat{m}')\,d\hat{m}' / \int_{\hat{m}}^{\infty} N(\hat{m}')\,d\hat{m}'. \tag{10}$$

If we assume the common form for the seismicity-magnitude relation given by equation (7b), then $P_I(\hat{m})$ can be incorporated in $P_C(\hat{m})$

$$P_C(\hat{m}) = \int_{\hat{m}}^{\infty} e^{-\alpha'\hat{m}'} P_I(\hat{m}')\,d\hat{m}' / \int_{\hat{m}}^{\infty} e^{-\alpha'\hat{m}'}\,d\hat{m}'.$$

Now assuming $P_I(\hat{m})$ to be the normal distribution function given by (2)(strictly speaking, we have seen that this is valid only for stations, not networks) and integrating the denominator of the above equation,

we have:

$$P_C(\hat{m}) = \alpha' e^{\alpha' \hat{m}} \int_{\hat{m}}^{\infty} e^{-\alpha' \hat{m}'} \Phi\left(\frac{\hat{m}' - \mu'}{\sigma_n}\right) d\hat{m}'$$

A linear transformation of the variable of integration, $\hat{m}' = \hat{m}'' + \mu'$, results in

$$P_C(\hat{m}) = \alpha' e^{\alpha'(\hat{m}-\mu')} \int_{\hat{m}-\mu'}^{\infty} e^{-\alpha' \hat{m}''} \Phi\left(\frac{\hat{m}''}{\sigma_n}\right) d\hat{m}''.$$

We now make the substitution of the error function

$$P_C(\hat{m}) = \frac{\alpha'}{2} e^{\alpha'(\hat{m}-\mu')} \int_{\hat{m}-\mu'}^{\infty} e^{-\alpha' \hat{m}''} d\hat{m}'' + \int_{\hat{m}-\mu'}^{\infty} e^{-\alpha' \hat{m}''} \operatorname{erf} \frac{\hat{m}''}{2^{1/2}\sigma_n} \, d\hat{m}''. \tag{11}$$

The first integral in (11) is well known:

$$\int_{\hat{m}-\mu'}^{\infty} e^{-\alpha' \hat{m}''} d\hat{m}'' = \frac{1}{\alpha'} e^{-\alpha'(\hat{m}-\mu')}.$$

The second integral in (11) can be evaluated (e.g., Abramowitz and Stegun, 1964, p. 304) and reduces straightforwardly to:

$$\int_{\hat{m}-\mu'}^{\infty} e^{-\alpha' \hat{m}''} \operatorname{erf}\left(\frac{m''}{2^{1/2}\sigma_n}\right) d\hat{m}'' = \frac{1}{\alpha'} e^{\sigma_n^2 \alpha'^2/2} \left\{1 + \exp\left[-\alpha'(\hat{m}-\mu') - \frac{\sigma_n^2 \alpha'^2}{2}\right] \cdot \operatorname{erf}\left(\frac{\hat{m} - \mu'}{2^{1/2}\sigma_n}\right) - \operatorname{erf}\left(\frac{\hat{m} - \mu' + \alpha' \sigma_n^2}{2^{1/2}\sigma_n}\right)\right\}.$$

Thus, substituting these in (11) and performing some algebraic manipulation, we have:

$$P_C(\hat{m}) = \frac{1}{2}\left\{1 + \operatorname{erf}\left(\frac{\hat{m} - \mu'}{2^{1/2}\sigma_n}\right) + \exp\left[\alpha'(\hat{m} - \mu') + \frac{\sigma_n^2\alpha'^2}{2}\right] \cdot \left[1 - \operatorname{erf}\left(\frac{\hat{m} - \mu' + \alpha'\sigma_n^2}{2^{1/2}\sigma_n}\right)\right]\right\}.$$

Our assumption for $P_I(\hat{m})$ in equation (2) requires

$$P_I(\hat{m}) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\hat{m} - \mu'}{2^{1/2}\sigma_n}\right)\right].$$

Our result then for the probability bias in the cumulative detection curve versus the incremental is:

$$P_C(\hat{m}) - P_I(\hat{m}) = \frac{1}{2}\exp\left[\alpha'(\hat{m} - \mu') + \frac{\sigma_n^2\alpha'^2}{2}\right] \cdot \left[1 - \operatorname{erf}\left(\frac{\hat{m} - \mu' + \alpha'\sigma_n^2}{2^{1/2}\sigma_n}\right)\right]. \quad (12)$$

Lacoss (1972) has previously expressed this bias in a form equivalent to (12). We list in Table II the magnitude bias of the cumulative threshold determination relative to the incremental determination predicted by (12) at the 90%, 50% and 10% thresholds for various reasonable values of σ_n and α; and the bias is significant in all cases. The bias given by

(12) applies strictly to a single station since $P_I(\hat{m})$ for a network has the hump discussed in the previous section and cannot be expressed as a normal distribution function. However, we can again expect the bias to be toward lower magnitude, by some unpredictable amount.

Note that the bias given by (12) is for the cumulative curve relative to the incremental curve. If the 90% threshold obtained from the incremental curve is already biased low as discussed in the previous section, the cumulative threshold determination will be even further biased low. For example, if $\sigma_n = \sigma_s = .3$ and $\alpha = 1$ and the 90% threshold is considered, then from Table II the cumulative threshold will be .21 lower than the incremental threshold which is already biased .16 lower than the threshold based on operational magnitude in this case. This is a total bias of -.37 magnitude unit, a significant amount indeed. Since there is no incremental bias for the 50% threshold, the cumulative bias only need be considered; for our example it amounts to -.34 from Table II, again a significant amount.

We have now derived the errors to be expected when determining the threshold magnitude of a station using magnitudes calculated at that station, and have indicated what errors to expect when networks are considered rather than a single station. This knowledge allows us to remove the errors and make a valid comparison between empirical thresholds and theoretical

ones based on operational magnitude. What corrections are necessary if we determine empirically station or network thresholds relative to an independent magnitude scale (e.g., surface-wave threshold on m_b) and then wish to compare them to theoretical ones based on operational magnitude? This is the topic of the next section.

CONVERTING DETECTION THRESHOLDS FROM M_s to m_b

Figure 2 shows a hypothetical but typical group of data with M_s's from a network whose threshold we wish to determine and m_b's from the same or an independent network. By plotting the percentage of events at given m_b's for which a surface wave was detected, one can simply determine a probability of detection curve for surface waves based on event m_b. The problem is to convert a surface-wave threshold on m_b which was determined directly using the given m_b's to an equivalent threshold on M_s, or vice-versa. The simplest approach in transforming LR thresholds from the M_s magnitude scale to the m_b scale is to equate detection probabilities for corresponding points on the M_s-m_b line which fits a set of empirical data from the network or station under consideration. This approach, however, leads to prediction of LR thresholds on m_b which are too low when compared to those actually determined on m_b. The problem is indeed more complex, and one must take into account the distributional properties of the data, such as M_s and m_b scatter about the M_s-m_b line and increasing numbers of events with decreasing magnitude, as shown in Figure 2. The correct transformation of thresholds from the M_s to m_b scales and vice-versa will now be presented.

The detection probability of surface waves conditional upon m_b can be expressed in terms of that conditional upon M_s by:

$$P(m_b) = \int_{-\infty}^{\infty} f(M_s \mid m_b) \cdot P(M_s) \cdot dM_s \qquad (13)$$

This is a useful relation; that is, we can convert predicted thresholds for surface waves in terms of M_s into terms of m_b to compare with empirical thresholds determined directly. To evaluate (13) in practice, we need explicit expressions for the two factors in the integrand. The second factor is given for a single station by equation (1). (For a network, (1) may be a satisfactory approximation, since the detection probability curve for a network is an F distribution (Herrin and Tucker, 1972) which approaches a normal distribution as the number of stations in the network increases.) The first factor in (13) can be expressed by the basic probability identity:

$$f(M_s \mid m_b) = f(m_b, M_s)/f(m_b) \qquad (14)$$

The joint density function $f(m_b, M_s)$ of the operational magnitudes reflects the scatter in a typical M_s versus m_b plot; this density function is expressed by introducing the true magnitudes, $\tilde{M}_s$ and $\tilde{m}_b$, each pair of which are linearly related in a precise manner thus:

$$\tilde{M}_s = a\tilde{m}_b + b. \qquad (15)$$

We assume that M_s and m_b are normal random variables with means $\tilde{M}_s$ and $\tilde{m}_b$ and standard deviations σ_{ms} and

σ_{mb}. Then the joint density function of the complete population of events arising from events with true magnitude $\tilde{m}_b$ is:

$$f(m_b, M_s) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_{ms}\sigma_{mb}} e^{-Q(m_b, M_s)} f(\tilde{m}_b) \cdot d\tilde{m}_b . \qquad (16)$$

The exponent Q in the above is

$$Q(m_b, M_s) = \frac{1}{2}\left[\frac{(m_b - \tilde{m}_b)^2}{\sigma_{mb}^2} + \frac{(M_s - \tilde{M}_s)^2}{\sigma_{ms}^2}\right] . \qquad (17)$$

$\tilde{M}_s$ could just as well serve as the variable of integration in (16). Now let the form of the density function of m_b be given by equation (7a). Then it is apparent from a comparison of equations (7a) and (7b) that the density function for $\tilde{m}_b$ is:

$$f(\tilde{m}_b) = \beta_b' e^{-\alpha' b \tilde{m}_b} e^{-\alpha'^2_b \sigma^2_{mb}/2} \qquad (18)$$

Note that $f(\tilde{m}_b)$ will be equivalent to $f(m_b)$ only if $\sigma_{mb}=0$ or $\alpha=0$. This $f(\tilde{m}_b)$ is not a proper probability density expression in that it, as well as $f(m_b)$ and thus $f(m_b, M_s)$ above, does not integrate to unity; however, the result we need, $f(M_s | m_b)$, is a ratio of these as given by (14) and is finite and integrable to unity. Using (15) in (17) and then substituting (17) and (18) in (16) and performing the integration results in

$$f(m_b, M_s) = \frac{\beta' _b \exp \frac{-\alpha'^2_b \sigma^2_{mb}}{2}}{(2\pi)^{1/2}(\sigma^2_{ms}+a^2\sigma^2_{mb})^{1/2}} \exp \left\{ -\frac{1}{2}\left[\frac{[M_s - (am_b + b)]^2}{\sigma^2_{ms} + a^2\sigma^2_{mb}}\right] \right.$$

$$\left. + \frac{1}{2}\left[\frac{-2\alpha'_b(\sigma^2_{ms} m_b + a\sigma^2_{mb}(M_s - b)) + \alpha'^2_b \sigma^2_{ms}\sigma^2_{mb}}{\sigma^2_{ms} + a^2\sigma^2_{mb}}\right]\right\} \qquad (19)$$

Using (7a) and (19) in (14) we arrive at

$$f(M_s|m_b) = \frac{1}{\sqrt{2\pi}(\sigma^2_{ms} + a^2\sigma^2_{mb})^{1/2}} \exp \left\{ -\frac{1}{2(\sigma^2_{ms} + a^2\sigma^2_{mb})} \right.$$

$$\left. [M_s - (am_b + b - \alpha'_b a\sigma^2_{mb})]^2 \right\} \qquad (20)$$

Note that this is a normal probability density function with a mean $\alpha'_b a\sigma^2_{mb}$ units below the true M_s-m_b line and a variance $\sigma^2_{ms} + a^2\sigma^2_{mb}$. Equation (20) should accurately describe the density of points along any vertical line in an M_s-m_b plot such as Figure 2, provided that the LR phase from all events with the given m_b is detected by the station or network. To illustrate (20) we take a simple, but realistic, case: a=1 and b=0 (M_s=m_b), $N \propto e^{-\alpha' b m_b}$ (natural seismicity), and σ_{ms}=σ_{mb} (equivalent M_s and m_b scatter). The conditional distribution of M_s is then:

$$f(M_s \mid m_b) = \frac{1}{(2\pi)^{1/2} 2^{1/2} \sigma_{ms}} \exp \left\{ -\frac{1}{2} \left[\frac{M_s - (m_b - \alpha' b \sigma^2_{ms})}{2^{1/2} \sigma_{ms}} \right]^2 \right\}. \quad (21)$$

So, with $M_s = m_b$, we have a normal density function for M_s with mean $m_b - \alpha' b \sigma^2_{ms}$ and variance $2\sigma^2_{ms}$; the mean of the distribution lies $\alpha' b \sigma^2_{ms}$ below the true $M_s - m_b$ line. Taking the above case again, but now allowing $\sigma_{mb} = 0$ (no scatter in m_b) and $\alpha' = 0$ (uniform seismicity), we get:

$$f(M_s \mid m_b) = \frac{1}{(2\pi)^{1/2} \sigma_{ms}} \exp \left[-\frac{1}{2} \left(\frac{M_s - m_b}{\sigma_{ms}} \right)^2 \right]. \quad (22)$$

Here we have simply a normal density function for M_s, with mean m_b and variance σ^2_{ms}. This is the density function proposed by Harley and Heiting (1972) and Lacoss (1971); however, (21) is the proper expression, and it includes a downward displacement of the mean due to increasing seismicity with decreasing magnitude.

We now know both factors needed in (13) to perform the conversion of surface-wave thresholds from M_s to m_b scales. However, we note that in dealing with observed magnitudes we need $P(\hat{m}_b)$ rather than $P(m_b)$. If we assume that $\hat{m}_b$ scatters normally about

m_b with variance σ_{sb}, the form of the above expression is preserved with $(\sigma_{mb}^2 + \sigma_{sb}^2)^{1/2}$ replacing σ_{mb} everywhere. The result of using m_b rather than $\hat{m}_b$ is to predict thresholds by (13) which lie at a lower magnitude than observed; that is, any probability calculated by (13) for m_b really applies to a higher $\hat{m}_b$. If the $\hat{m}_b$ is a network average, the overall effect is small, less than .1 magnitude unit in the threshold value. A similar consideration of $\hat{M}_s$ rather than M_s would provide the necessary distribution $f(\hat{M}_s, \hat{m}_b)$ if one desired to use $\hat{M}_s$ threshold calculations given by (2).

It is now possible, using (20) and (1), to evaluate (13), and it is seen that this evaluation requires knowledge or estimation of several parameters. In our application of (13), we therefore estimate an empirical M_s-m_b line, correct the M_s-m_b line for the Herrin and Tucker bias at low magnitudes (as discussed in a later section), choose a seismicity-magnitude relation from some independent source, and assume a normal distribution for the magnitude scatter, the variance of which depends on source bias and path effects.

A biased estimate of $\hat{M}_s$ by the station or network due to favorable or unfavorable travel paths for LR affects the surface-wave detection threshold on the m_b scale, calculated by (13), through the empirical M_s-m_b relation. Empirical surface-wave thresholds on $\hat{m}_b$ will be equally affected though, and so this $\hat{M}_s$ bias need not be considered when comparing calculated with empirical detection thresholds on m_b. Similarly, any

bias in $\hat{m}_b$ values can be ignored in converting surface-wave thresholds from the M_s to m_b scales, because the empirical M_s-m_b scale again compensates for this bias. However, in reference to other stations or networks, the m_b bias cannot be ignored. The reason is that, if one were to make comparisons of LR thresholds on $\hat{m}_b$ between two stations or networks whose sources of m_b values were not identical, then discrepancies could arise if one set of $\hat{m}_b$'s were biased relative to the other. These discrepancies can only be eliminated by estimating and removing the relative bias in observed $\hat{m}_b$.

The calculated detection curves in Figure 3 show that the effect in (13) of using α=1.0 (equation 21) is to raise the calculated 90% threshold on the m_b scale by an additional .1 to .2 magnitude unit over the case where α=0.0 (equation 22) when a normal distribution with $\sigma_{ms}=\sigma_{mb}=.3$ is assumed for m_b. The remainder of the detection curve is also shifted to higher m_b, though not by equal amounts. This is in agreement with simulation results to be given later.

One may want to perform the inverse of the above problem; that is, one may have determined empirically surface wave detection probabilities on the m_b scale and desire to transform them to the M_s scale. The apparent (but incorrect) procedure is simply to substitute M_s for m_b and vice versa in equation (13) and

proceed with the numerical integration as before, only now to obtain detection probabilities at the incremental M_s's. This procedure is invalid since the empirical probability of detecting a surface wave from an event of magnitude m_b is a result of looking for the surface waves from events of varying M_s at that m_b, not just a particular M_s. Therefore one must assume some detection probability curve as a function of M_s in order to evaluate $P(m_b)$ which would be within the integral when M_s and m_b were interchanged; but this detection probability curve $P(M_s)$ is exactly what we are attempting to determine.

We conclude that there is no direct way of going from an m_b to a M_s scale for surface wave detection probabilities. This suggests than an iterative algorithm should be used to arrive at the correct detection probability curve on the M_s scale. To do this we take a starting $P(M_s)$ curve and perturb its mean (50% point) and variance while making the transformation according to equation (13). This requires that an M_s-m_b relationship for the network or station under consideration, a seismicity-magnitude relationship, and a standard deviation for m_b and M_s error be chosen in order to calculate the factors in (13). The calculated curve on the m_b scale $P(m_b)$ which best approximates the empirical one indicates the correct M_s detection probability curve, $P(M_s)$. Figure 4 gives the results of a program which performs this algorithm, in this case for M_s=1.0 m_b, α=1.0, and $\sigma_{mb}=\sigma_{ms}$=0.25 for body-wave magnitude

and surface wave magnitudes, probably a conservative value. In the same figure the curve resulting when $M_s = 1.2\ m_b - 1.0$ is used demonstrates the sensitivity of the spread of the calculated $P(M_s)$ curve to the slope of the M_s-m_b relation. The resulting $P(M_s)$ curve is also dependent on the intercept of the M_s-m_b relation and would slide along the horizontal axis according to the exact value of this intercept. We also show in Figure 4 the result when $\alpha = 0.0$ is used; evidently the threshold magnitudes on M_s would be significantly overestimated if seismicity constant with magnitude were assumed when the true relation had $\alpha = 1.0$.

This section has provided a procedure for transforming thresholds from one magnitude scale to another, using LR thresholds on m_b and M_s as the example. Again, the results are strictly for a single station; results for some networks may be closely approximated by those for a single-station though. The importance of the seismicity-magnitude relation in the threshold theory was again demonstrated; and in this case of transforming thresholds between magnitude scales, the importance of the functional relation between magnitudes on the two scales became evident.

DETERMINING THE TRUE M_s-m_b LINE FROM OBSERVED DATA

One of the important problems in seismic discrimination is to estimate the M_s-m_b relation for a given group of observed (M_s, m_b) pairs. Determination of the true relation at or below the threshold for surface-waves (usually found to be higher than the body-wave threshold) is complicated by an upward M_s bias in the observed data points for this range as discussed in Herrin and Tucker (1972) and earlier in this report. It is possible, however, in theory to reconstruct the true M_s-m_b relation from the observed data although there will be difficulties in practice. The procedure is as follows:

For simplicity, let $M_s = m_b$, $\sigma_{ms} = \sigma_{mb}$, and $\sigma_{ss} = \sigma_{sb}$. Results for the general case where parameters vary can be apparent from the results we will attain with these simplifications. Equation (21) gave the density function $f(M_s | m_b)$ for a single station; for $f(\hat{M}_s | \hat{m}_b)$ appropriate to observed magnitudes, we add the effects of signal variance to get:

$$f(\hat{M}_s | \hat{m}_b) = \frac{1}{(2\pi)^{1/2}(2)^{1/2}(\sigma_{ms}^2 + \sigma_{ss}^2)^{1/2}} \exp\left(-\frac{1}{2}\left\{\frac{\hat{M}_s - \left[\hat{m}_b - \alpha_b'(\sigma_{ms}^2 + \sigma_{ss}^2)\right]}{(2)^{1/2}(\sigma_{ms}^2 + \sigma_{ss}^2)^{1/2}}\right\}^2\right).$$

The mean of this normal density function parallels the true M_s-m_b line, and the variance is constant over all $\hat{m}_b$. Then, for a given ε, $0 < \varepsilon < 1$, the integral of the density function at a given $\hat{m}_{b_i}$:

$$\varepsilon = \int_{\hat{M}_{s_i}}^{\infty} f(\hat{M}_s \mid \hat{m}_{b_i})\, d\hat{M}_s \tag{23}$$

must be satisfied with an $\hat{M}_{s_i}$ such that all $(\hat{M}_{s_i}, \hat{m}_{b_i})$ pairs will parallel the true M_s-m_b line. One can specify exactly that ε which will cause $\hat{M}_{s_i}$ to fall on the true M_s-m_b line. In the case $M_s = m_b$, let the integral (23) be evaluated to $\hat{M}_{s_i} = \hat{m}_{b_i}$

$$\varepsilon = \int_{\hat{m}_{b_i}}^{\infty} Z \left\{ \frac{\hat{M}_s - \left[\hat{m}_{b_i} - \alpha_b' (\sigma_{ms}^2 + \sigma_{ss}^2) \right]}{(2)^{1/2} (\sigma_{ms}^2 + \sigma_{ss}^2)^{1/2}} \right\} d\hat{M}_s .$$

Using the properties of the normal density function, this can be transformed to an integral over the unit normal density:

$$\varepsilon = \int_{\frac{\alpha_b' (\sigma_{ms}^2 + \sigma_{ss}^2)^{1/2}}{(2)^{1/2}}}^{\infty} Z(z)\, dz .$$

(This result would be unchanged if $M_s = am_b + b$.) For example, if $\alpha_b = 1$ and $\sigma_{ms} = \sigma_{ss} = .212$, then $\varepsilon \approx .32$

would be required to have $\hat{M}_{s_i}$ on the M_s-m_b line; repetition of this at various $\hat{m}_{b_i}$ would define the true M_s-m_b line by a series of points. This procedure could be applied only in the ideal case where both surface and body waves were detected at the station for all events.

With real data, one must account for thresholds; that is, all the points predicted by the above density function will not appear in the data due to nondetections. The observed density of points is related to $f(\hat{M}_s|\hat{m}_{b_i})$ by

$$f(\hat{M}_s|\hat{m}_{b_i}) = f(\hat{M}_s^o\ \hat{m}_{b_i}^o)/[P_{sw}(\hat{M}_s) \cdot P_{bw}(\hat{m}_{b_i})]$$

where P_{sw} and P_{bw} are the detection probabilities for surface and body waves in the form given by equation (2) earlier. By definition here, $f(\hat{M}_s^o|\hat{m}_{b_i}^o) \leq 1$. Using this reduced density function for observed data and substituting in (23), we have then:

$$\varepsilon = \int_{\hat{M}_{s_i}}^{\infty} \left\{ f(\hat{M}_s^o|\hat{m}_{b_i}^o)/[P_{sw}(\hat{M}_s) \cdot P_{bw}(\hat{m}_{b_i})] \right\} d\hat{M}_s. \qquad (24)$$

Using an actual plot of observed $(\hat{M}_s, \hat{m}_b)$ pairs, one can evaluate (24) numerically down to the $\hat{M}_{s_i}$ at which the integral equals ε. In order to do this, the seismicity-magnitude relation $\log N = -\alpha_b \hat{m}_b + \beta$ and

P_{sw} and P_{bw} must be estimated by standard procedures. From the seismicity-magnitude relation, the number of events expected to lie at each $\hat{m}_{b_i}$, say N_i, can be determined; and if we let N_j be the number of observed events in each increment of $\hat{M}_s$ at that $\hat{m}_{b_i}$, we can approximate equation (24) by:

$$\varepsilon = \sum_{\hat{M}_{s_j}=\hat{M}_{s_{max}}}^{\hat{M}_{s_i}} \left(\frac{N_j}{N_i}\right) / [P_{sw}(\hat{M}_{s_j}) \cdot P_{bw}(\hat{m}_{b_i})]$$

or, removing the constants from the summation,

$$N_i P_{bw}(\hat{m}_{b_i})\varepsilon = \sum_{\hat{M}_{s_j}=\hat{M}_{s_{max}}}^{\hat{M}_{s_i}} N_j / P_{sw}(\hat{M}_{s_j}) \qquad (25)$$

where $\hat{M}_{s_{max}}$ is the largest $\hat{M}_s$ seen for the given $\hat{m}_{b_i}$. A choice of ε can still be made as shown above to make computed M_{s_i}'s lie on the true M_s-m_b line. But in most cases it would be preferable to use a smaller ε, say 0.1 or 0.2, so that the method can be applied down to a lower m_b; for with ε larger, one simply runs out of events due to nondetection of surface waves before the summation approaches $N_i P_{bw}(\hat{m}_{b_i})\varepsilon$.

An example of the application of equation (25) will be shown in the section on simulation. The procedure suggested here is valid only for single-station M_s-m_b plots. It requires rather precise

estimation of the probability detection curves of the station for body and surface waves; and if such precision is unattainable, then one should not extend the procedure below roughly the expected 90% threshold for either surface waves or body waves.

Before proceeding to a simulation approach to examine network threshold problems, we will treat in the next section one additional effect on threshold whose simple aspects can be handled analytically - the matter of correlated signals and noise among stations.

CORRELATED SIGNALS AND NOISE

Noise background at several stations of a network cannot be assumed to fluctuate independently because of seismic events or swarms of events which, considered as noise, can produce temporarily higher background at two or more stations and because of microseismic origins which can simultaneously affect many stations within range of these disturbances. In applying equation (1) to a network (Wirth, 1970) we must account for correlation, if any, among the noise amplitudes for the stations in the network. The effect of correlation in general depends on the magnitude distribution of the population of events under consideration.

If due to source bias a set of events have operational magnitudes varying about their true magnitude, then for this set of events there will be a correlation between the magnitudes observed at different stations. In other words, source bias implies signal correlation in reference to true magnitude. If the noise is also correlated as described above, it is of interest to determine the correlation between stations of the arguments of expression (1). The mathematical expressions of these relationships will be useful in assessing the effects of correlation on thresholds. We define the observed magnitudes $\hat{m}_1$ and $\hat{m}_2$ at two stations for events of true magnitude $\tilde{m}$ as the sum of two normal processes:

$$\hat{m}_1 = G(\tilde{m},\sigma_m) + G_1(0,\sigma_s)$$

$$\hat{m}_2 = G(\tilde{m},\sigma_m) + G_2(0,\sigma_s)$$

Here G is the random normal variate representing scatter of operational magnitude about the true magnitude, and G_1 and G_2 are random normal variates representing path effects which, when added to the operational magnitude, produce observed magnitudes. The correlation $\rho_{\hat{m}}$ between m_1 and m_2 is given by:

$$\rho_{\hat{m}} = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_s^2} \tag{26}$$

Note that $\rho_{\hat{m}}=0.5$ if $\sigma_s=\sigma_m\neq 0$ and $\rho_{\hat{m}}=0$ if $\sigma_m=0$. If $\hat{m}_1$ and $\hat{m}_2$ are normal processes with correlation as in (26), mean $\tilde{m}$, and standard deviation $(\sigma_m^2 + \sigma_s^2)^{1/2}$ and if μ'_1 and μ'_2 are zero-mean normal processes with parameters ρ_n and σ_n; the correlation ρ between $(\hat{m}_1-\mu'_1)$ and $(\hat{m}_2-\mu'_2)$ is given by:

$$\rho = \frac{\rho_{\hat{m}}(\sigma_m^2 + \sigma_s^2) + \rho_n\sigma_n^2}{\sigma_m^2 + \sigma_s^2 + \sigma_n^2} \tag{27}$$

Note that $\rho=\rho_{\hat{m}}=\rho_n$ if $\rho_{\hat{m}}$ and ρ_n are equal.

We can obtain analytical predictions for the effect of correlation on a network threshold in only the simplest of cases. A two-station network with both stations having an equal probability of detection of events can be treated through the bivariate normal distribution. Consider the case of random normal noise and signals so that the correlation of the variables $\hat{m}_1-\mu'_1$ and $\hat{m}_2-\mu'_2$ is as given in equation (27). Bivariate normal curves for ρ= 0.0, 0.5 and 1.0 were used to obtain the "one or more" detection probabilities for this hypothetical two-station network based on μ'=5.00, $(\sigma_m^2 + \sigma_s^2)^{1/2}$=0.0 and σ_n=0.2 as shown in Figure 5. The effect of perfectly correlated noise is to raise the 50% and 90% thresholds by little more than 0.1 magnitude unit.

We cannot predict analytically the noise correlation effect for realistic networks, though, and a simulation experiment is required. A program was written to generate populations of signals and noise at the various stations of a network according to the desired statistical parameters and to select out those "events" which were detected. This program will be described in more detail in a later section. To check the program, we compared its detection curve in the "one or more" detection case for a two-station network with μ'=5.00, σ_n=0.2, and ρ=.95 for the noise and $\sigma_m^2=\sigma_s^2=0$ for the signals to the theoretical prediction as shown in Figure 5. (The correlation coefficient in the simulation program was chosen to be .95 since exactly

1.0 would give negative eigenvalues for the noise matrix.) The simulated and predicted curves seem to agree well, keeping in mind this slight difference in the value of the noise correlation. We proceeded next to simulate a "two or more" detection curve for the nine-station Long Period Experimental network. The noise amplitudes and noise correlation coefficients were taken from von Seggern (1974); the ρ_n averaged roughly 0.4. An epicenter at 50N and 120E was arbitrarily chosen, and the simulation was run for σ_n=.2 and σ_s=.3, σ_m=0.0, and $\rho_{\hat{m}}$=0.0. The resulting detection curve as a function of true magnitude is shown in Figure 6 along with the one theoretically predicted for ρ=0. Correlated noise has apparently little or no effect in this case.

All the foregoing discussion on effects of correlated noise expresses thresholds in terms of true magnitude and assumes that the number of events versus magnitude is uniform. The true seismicity-magnitude relation is commonly approximated by $\log_{10}N=\alpha m_b+\beta$ though, and we must account for this in our simulation experiment. We will include such a case with correlated noise on a ten-station network in the next section, which will describe our simulation procedure and results.

SIMULATION EXPERIMENT

Need for a Simulation Approach

While we have been able in the previous sections to derive analytically many useful relationships between the thresholds and properties of a single station, or in some cases, a two-station network, many results required for practical evaluation of a real network could not be obtained. For example:

1. A network does not have a response similar to a single station, and analysis of the network response becomes prohibitively complicated when the different stations of the network have noise levels with differing means and variances and are at different distances from the epicentral area.

2. In a real network the P threshold influences the LR threshold by virtue of the fact that one does not attempt to detect LR unless P has been detected. While this will not raise the LR threshold if the P threshold is far below the LR threshold, many real networks do not have this property. For example, detection of LR by the LPE network in response to NOS epicenters does not (von Seggern, 1974), and a network of LASA stations would not. We shall show for this latter case that the thresholds are affected to the degree of 0.1-0.2 magnitude units.

3. Source bias due, for example, to differing fault mechanisms of an earthquake or differing depths of burial of an explosion influences thresholds.

4. Correlation of mean-square noise amplitudes among stations influences thresholds; but as we shall see in the simulations the effect is almost negligible for realistic correlation levels when source bias also exists.

5. Incremental and cumulative histograms of the number of events as a function of $\hat{m}_b$ for which LR has been detected are often of interest. However, analysis of such curves has thus far eluded analytical investigation even for single stations.

6. Determination of the true M_s-m_b relationship by fitting straight lines to observed data sets is unreliable due to magnitude bias introduced around the station and network thresholds. This bias can be evaluated, and thus corrected, by means of simulation.

NETWORTH Predictions

As we shall see below, it has been possible to use the standard program NETWORTH (Wirth, 1970) to compute thresholds for true magnitudes if there is no source bias or noise correlation, and to compute thresholds with respect to operational magnitudes if there is no noise correlation. Some modifications to NETWORTH are required in order to compute thresholds when noise correlation is present or to compute thresholds with respect to true magnitudes when signal correlation (source bias) is present. The modified program, NETWCORR, does not compute thresholds with respect

to observed magnitudes. Only the full simulation program MSBNET, to be discussed below, will accomplish that.

A detailed discussion of the new NETWCORR techniques is contained in Appendix II; however, a short outline will be given here. The user of NETWCORR now has the option of reading in a signal correlation matrix and a noise correlation matrix. Equation (26) may be used to calculate the elements of the signal correlation matrix from a given variance σ_m^2 of source bias and given variance σ_s^2 of signal amplitudes for an event. At each magnitude the technique outlined by Shumway and Blandford (1970) is then used to generate sets of random signal and noise amplitudes with the given correlation structures. Each set of such amplitudes represents an event and supplies arguments, whose correlation coefficient is given by equation (27), in equation (1) for all stations of the network. If this argument is greater than zero a detection by that station is declared. If k or more stations detect the event, then this increases the probability of a "k or more" network detection for an event of the magnitude under consideration. A sufficient number of random events is generated at each magnitude to produce a stable probability estimate.

Table III describes the only network considered in detail in this report. It has been designed to fit as closely as possible ten LASA-equivalent stations, each of which is 65° from the Kamchatka-Kuril region.

The μ_b' and μ_s' values are the estimated 50% threshold magnitudes for LASA, as will be discussed in a later section. The estimates for the signal variances are consistent with the work of von Seggern (1972) and Ericsson (1971a), and σ_n is consistent with short-period values for LASA beams obtained by Chang (1973). The same σ_n was assumed for long-period beams. The α value of 0.8 is roughly correct for the Kamchatka-Kuril region earthquakes on the m_b scale (W. Dean and R. Ahner, personal communication). We have assumed $M_s=1.0\ m_b$. A more reasonable formula is $M_s=1.0\ m_b-0.8$, calculated from a LASA-Kuriles M_s-m_b diagram in Mack (1971). Then if an observed M_s threshold is 3.00, one would translate that into 3.80 for comparison with this simulation experiment. The reverse procedure would need to be carried out to translate simulation M_s thresholds downward for comparison with observation.

In this report we consider only networks of type I in which a station does not attempt to detect LR unless it has detected P. A network of type II would be one in which all stations look for LR if a specified number of them detect P. The program MSBNET may be run for either type of network. For the examples considered in this report the changes in 90% thresholds are less than 0.1 between runs for the two types.

In Table IV we have the NETWORTH and NETWCORR results. Figure 7 gives complete curves for the standard NETWORTH results. The 90% points listed in

Table IV show that there is approximately 0.2 magnitude difference in the thresholds between the standard NETWORTH results and the NETWCORR results with source bias (with and without noise correlation) if the true magnitudes are used as reference. If, however, the operational magitudes are used as reference the difference is closer to 0.1 magnitude units. In both cases the predicted thresholds with source bias are higher than without. The differences in thresholds are generally smaller but still have the same sign for the 50% and 10% levels. Note that 50% true and operational thresholds for the case of source bias present are nearly the same, with or without noise correlation.

As far as we know these NETWORTH results are correct if the underlying reasonable assumptions are correct. This does not, however, mean that they will agree with observation because observed magnitudes are different from both operational and true magnitudes. Thus NETWORTH may perfectly predict the probability of detecting an explosion with a specified yield (translated into true magnitude) and yet not be verified by observation if observed magnitudes (instead of yields) are used to determine the threshold. The only way in practice to establish the verifying relationships is by means of the simulation techniques to be discussed in the next section.

MSBNET Simulation Techniques

A complete discussion of program MSBNET is given in Appendix III; here we give only a short outline. The user specifies station parameters, means and standard deviations of the noise at the stations, the standard deviation of the signal, the seismicity curve together with the magnitude limit below which true magnitude events are not to be randomly generated, the portion of the signal variance due to source bias, the noise correlation matrix, and the magnitude limits between which M_s-m_b regressions are to be calculated.

Let us first consider the case where there is no interstation noise correlation. Consider the lowest 0.1 true magnitude increment in which random events are generated. For a typical run it will comprise $\sim 10^3$ "events" with the appropriate true $\tilde{m}_b$ value and the corresponding true $\tilde{M}_s$ value. Consider now the first network station. To each of the k^{th} pair of true magnitudes are added two random normal numbers with zero mean and the appropriate standard deviations. One of these numbers is the source bias and for each individual event will be the same at all stations. The other number is the "unexplained" signal variance due to travel path effects, and this will be different for each station and each event. Signal amplitude correlation in this situation may be calculated by equation (26). The argument of equation (2), and thus the probability of detection, may then be calculated

for both P and LR using the newly generated observed $\hat{m}_b$ and $\hat{M}_s$ values. Two uniform random numbers are then generated in the interval (0,1); and if these numbers are less than the corresponding probabilities of detection, the phase is "detected" at that station.

To make the following discussion clear it seems necessary to be somewhat abstract. For ease of reference we choose to use the same variable names as are used in program MSBNET. The variable names are suggestive of their role in the program.

Suppose the P wave from event K is detected by a station; then NOMBD(K) is increased by 1, and the observed value of m_b is added to AMBM(K). The same procedure is followed for LR and $\hat{M}_s$, using NOMSD(K) and AMSM(K) if the P wave is detected for network type I, or in either case for type II. After all stations for all events for all magnitude increments have been considered, if NOMBD(K)$\geq$NRMBD then AMBN(K)/NOMBD(K)=$\bar{\hat{m}}_b$ is computed. (NRMBD is the number of P-wave detections required for a network detection; if NOMBD(K)<NRMBD we go to event K+1.) Then the counter in the NMBNTD appropriate to $\bar{\hat{m}}_b$ is increased. Then if NOMSD(K)$\geq$NRMSD, we compute AMSM(K)/NOMSD(K)=$\bar{\hat{M}}_s$ and increase the counter in the appropriate NMSNTD magnitude cell <u>and</u> in the NMBMSD magnitude cell determined by $\bar{\hat{m}}_b$. (But if NOMSD(K)<NRMSD, the number of LR detections required for a network detection, we go to event K+1.) After all events are considered, regression curves may be computed on the surviving $\bar{\hat{m}}_b$, $\bar{\hat{M}}_s$ pairs.

The entire procedure may then be repeated for statistical stability, starting with a new random number and accumulating more events in the NMBNTD, NMSNTD, and NMBMSD arrays. Similar statistics are kept for the individual stations. Finally, the numbers required for the incremental, cumulative, and direct plots to be discussed in the next section may be output.

If one introduces a non-zero noise correlation matrix, only one small change is required in the main program. Due to core storage limitations an auxiliary program is used to store on tape several thousand sets of N numbers (where N is the number of stations in the simulated network) which have the means, standard deviations, and correlation matrix required for the noise. Then for each station the difference is taken between the randomly generated signal value and the randomly generated noise value, read from tape. If the signal is greater than the noise, a detection is declared.

Simulation Results

First let us discuss a single station trying to detect all P (or LR) signals. Given the parameters in Table III, an answer for P may be converted into an answer for LR simply by shifting a "P" graph 3.80-3.47=0.33 units (the M_s-m_b threshold difference at LASA for Kuril-Kamchatka) to the right. Rows 3 and 4 of Table V give us these results, and they are derived from Figure 8 as indicated in the third column.

We note that the 50% incremental value is in exact agreement with the mean noise levels in Table III. This is as predicted in the theoretical section of this report. Also, the .08 m_b bias in the asymptotic portion of the observed seismicity curve is in agreement with theory. In fact both curves, cumulative and incremental, are in complete agreement with theory. Note that for a single station, source bias cannot be distinguished from other sources of signal variation and that the correlation of noise has no effect on threshold. We note that the 90% observed incremental threshold is biased by 0.22 with respect to the true magnitude threshold in Table IV and by 0.12 with respect to the operational magnitude threshold.

Next we consider a single station which must first detect on P before it can detect on LR (we note that this is also a good model for two separate short-period arrays, e.g., NORSAR and LASA, one of which is prompted by the other). Figure 9 illustrates the LR thresholds on observed $\hat{M}_s$. For the parameters given here, comparing rows 4 and 7 in Table V, we see that there is very little effect on the observed $\hat{M}_s$ thresholds when the P threshold lies near that of LR. This mode of detection, however, offers the possibility of other kinds of plots, such as those shown in Figures 10 and 11. In Figure 10 we give incremental and cumulative plots of the number of events, as a function of observed $\hat{m}_b$, for which LR was detected. In Figure 11 we plot the fraction of events for given observed $\hat{m}_b$ for which

LR was detected. Various thresholds in terms of observed $\hat{m}_b$ for detection of LR may be obtained from these figures and are given in row 6 of Table V. Comparing rows 6 and 7 we see that the $\hat{m}_b$ 90% thresholds are much greater than the $\hat{M}_s$ thresholds, although the 50% points are rather close. At the 90% level the closest agreement of all may be found between the cumulative $\hat{m}_b$ and incremental NETWORTH thresholds for true magnitude $\tilde{M}_s$, 4.25 and 4.26 respectively. This agreement would, of course, be very difficult to predict from physical insight (we must remember that if $M_s=m_b-0.8$ instead of $M_s=m_b$, then the 4.26 value would become 3.46, but the predictions would still be understood to be in agreement).

We might ask what would be the effect of an even lower P threshold. Comparison of rows 5 and 6 of Table V shows that for a P station threshold of 3.00 instead of 3.47 there is almost no change in the incremental 50% and 90% LR thresholds on m_b. As we can see by comparing Figures 12 and 10, however, the 10% incremental threshold would be significantly changed. It is interesting that in varying the P threshold over a large range, even up to values greater than 3.80 (the LR threshold), the single station direct probability curve in Figure 11 did not change noticeably. This result is reflected in Table V. It would be of interest to prove this fact theoretically. We shall see below that this invariance does not hold with respect to <u>network</u> thresholds.

Again, these single-station results are unaffected by source bias or by noise correlation between stations. However, there would be an effect if there were correlation between the amplitude levels of short-period and long-period noise; this topic is not treated in this report.

Next we move to the subject of network thresholds. The thresholds predicted by NETWORTH have already been illustrated in Figure 7. We investigate first the problem of four stations out of ten detecting without prompting. Figure 13 is the appropriate plot. Although the horizontal axis is labeled $\hat{m}_b$, it could as well be $\hat{M}_S$ if all values were increased by 0.33 (assuming $M_S = m_b$).

We note, as discussed in the theoretical section, that there is a hump in the seismicity-magnitude curve in contrast to the single-station case. By comparison with Figure 23, for which no source bias was assumed, we see that the size of the hump is reduced by the assumption of source bias (Figures 22 and 23-31 correspond to 11 and 13-21 but are without source bias; Table VI corresponds to Table V). We illustrate in Figure 13 how it is possible to estimate erroneously the seismicity curve and to obtain α=1.2 instead of 0.8 in the seismicity-magnitude relation. Knowing that the hump is present, we can draw an accurate seismicity line using large events or we can estimate the slope of the seismicity lines from single-station incremental plots which have no such hump and thus

estimate thresholds which are reasonably close to the predicted ones. We find that the seismicity bias is 0.05 m_b, approximately what equation (8) would predict for a single station with a standard deviation of the signal equal to the source bias, 0.212. This is a very reasonable result in that for large magnitudes one would expect the average observed magnitudes to be very close to the true magnitudes except for source bias (Figure 23 indicates that there is no detectible seismicity-magnitude bias when source bias is absent).

Comparison of rows 11 and 12 from Table V shows that the 90% incremental threshold is in good agreement with the 90% true magnitude standard NETWORTH threshold; 3.51 as compared to 3.52. The preferred comparison would be with NETWORTH results including source bias, for either true or operational magnitude. From Table IV, these numbers would be 4.02 or 3.84 minus 0.33 (3.69 or 3.51) for true and operational magnitude, respectively, instead of 3.52; thus the agreement is quite good using operational magnitude, but not so good using true magnitude. Note that the network 4/10 50% point is in almost exact agreement with the single station 50% point, 3.46 compared to 3.47. This must, of course, be regarded as a coincidence in some sense since the 50% point is different for 1/10 or 2/10 station detection.

Finally let us discuss the complete network system where we require 1/10 LR detection following 4/10 P

detection. Figures 17a and 27, with and without source bias respectively, give the incremental and cumulative $\hat{M}_s$ curves. We note again that the hump is noticeably reduced by the presence of source bias. Figures 16a and 26 show that plotting the incremental and cumulative curves as a function of the $\hat{m}_b$ corresponding to the LR detection instead as of a function of the $\hat{M}_s$ value greatly reduces the hump and that when source bias is included the hump apparently disappears completely. The direct thresholds are determined from Figure 11. All the thresholds calculatd assuming source bias may be found in rows 15 and 16 of Table V. Rows 13 and 14 of both Table V and VI were calculted assuming that the individual station P threshold is lowered from 3.47 to 3.00. Comparison shows that the 90% thresholds are hardly changed but that the 50% levels change by as much as 0.21 magnitude units.

Due to the hump, the incremental 50% and 90% $\hat{M}_s$ thresholds are much closer together in the simulation results than in the NETWORTH results; no satisfactory fit with Table IV is possible. Also, it does not seem possible to fit any other of the simulation results closely to any of the NETWORTH results. Of course we should emphasize that there is no reason to expect a particularly good fit; the whole point of the simulation is to derive the behavior of thresholds obtained in specified empirical ways so that they may be related to the true or operational NETWORTH thresholds. In Table V, comparison of rows 15 and 16 to rows 21 and 22

from Figures 16b and 17b shows that if one has ρ_n=0.4 in addition to source bias, the effect is almost unnoticeable at the 50% level; and on the order of 0.05 magnitude units at the 90% level.

Similar results are obtained for 2/10 and 4/10 LR detection, and these results are given in Figures 18-21 when source bias is included. Compared to the 1/10 thresholds, the 2/10 and 4/10 thresholds increase by 0.10-0.15 with each step increase in number of detecting stations required. This is in rough agreement with, but slightly less than, the NETWORTH differentials of 0.15-0.20.

Figure 32 shows the probability of detection of LR by 1/10, 2/10 and 4/10 stations as a function of the observed $\hat{m}_b$ determined by four or more stations. The curves are plotted on log-probability paper on which the cumulative normal probability distribution function would plot as a straight line. The failure of the 2/10 and 4/10 curves to approximate a straight line suggests that fitting observed network curves with this function could on occasion be inappropriate. Note that this function fits the single-station data well though, as expected.

The main point which the reader should derive from a detailed study of the foregoing simulation results for a particular ideal network is that any casual approach to the estimation of seismic thresholds for a real network is likely to be in error by several tenths

of a magnitude unit. The best course of action when a real network is under consideration is to construct a model by estimating its parameters from the most accurate available data and to run a simulation using MSBNET on this model.

Estimating M_s-m_b Relationships

Finally we treat the question of the proper evaluation of M_s-m_b relations. Table VII gives the calculated slopes and intercepts using regression on $\hat{m}_b$, regression on $\hat{M}_s$, and maximum-likelihood (Ericsson, 1971b), for seven different types of M_s-m_b data cutoffs and for 1/10, 2/10, 4/10 LR detection. The data are the same as that from which the thresholds presented in Table V were estimated, and the simulation values for the slope and intercept are 1.0 and 0.0, respectively.

Column I of Table VII gives results of calculations which include basically all the network $\hat{M}_s$-$\hat{m}_b$ values; that is, those between magnitudes 3.0 and 6.0. Column II, using data from 4.0-6.0, may be regarded as high-magnitude estimates; and Column III, using data from 3.0-5.0 as low-magnitude estimates. Columns IV and V encompass the magnitude interval from 6.0 down to that $\hat{M}_s$ above which lie 90% or 80%, respectively, of the total number of detected events. Column VI cuts off the data at the $\hat{m}_b$ for which there is a 90% probability of detecting LR.

Columns I-VI have symmetrical cut-offs; that is, the $\hat{m}_b$ and $\hat{M}_s$ cut-offs are equal. This means that the

cutoff lines intersect at the mean $\hat{M}_s$-$\hat{m}_b$ line. For Columns VII and VIII, this is not true. In these columns there is the standard all-inclusive 3.0-6.0 range for the $\hat{m}_b$ values; however, for Column VII the lower cut-off on $\hat{M}_s$ is given by the $\hat{M}_s$ value equal to the 90% direct threshold on $\hat{m}_b$, and for Column VIII by the $\hat{M}_s$ value above which lie 90% of all events for which LR was detected.

Inspection of this table shows that accurate estimation of the parameters a and b are obtained using symmetrical cutoffs and maximum-likelihood estimation if both low-magnitude cutoffs are given by 4.0 or if a cutoff is made at the $\hat{M}_s$ value above which lie 90% or 80% of events with detected LR (Columns II, IV, V). Accurate results are also obtained by regression on $\hat{M}_s$ if there is no cutoff on $\hat{m}_b$ and if $\hat{M}_s$ cutoff is given by the $\hat{M}_s$ value above which lie 90% of all events with detected LR (Column VII) or is given by the $\hat{M}_s$ value equal to the LR 90% direct threshold (Column VIII). For none of the cases considered is regression on $\hat{m}_b$ accurate. However, in all of these cases there is some bias; and a simulation approach should be used to predict it and thus enable one to remove it.

The success of the maximum-likelihood approach is not surprising because after low magnitudes are removed and the remaining data is cut off so as to retain the symmetry of the data about the true M_s-m_b relation, then the maximum-likelihood approach would be expected

to be unbiased toward either too large or too small a slope.

The success of the regression on $\hat{M}_s$ is perhaps surprising. It certainly has nothing to do with greater reliability on $\hat{M}_s$ values. In fact, observed $\hat{M}_s$ values in the simulation will have larger variances than the observed $\hat{m}_b$ values when LR detection is not as good as P detection. However, such a cutoff will ensure that at each $\hat{M}_s$ value, $\hat{m}_b$ values will be scatter approximately normally about the mean line (this is not exactly true, since we have seen in the theoretical section that realistic seismicity shifts the mean of the normal distribution). On the other hand, the same cutoff will ensure that near this cutoff the $\hat{M}_s$ values are quite dramatically <u>not</u> normally distributed about the mean line, for fixed $\hat{m}_b$. And, of course, such a normal distribution is a requirement for validity of regression calculations.

The point of this section is that one should simulate the network under consideration and so discover by trial and error which techniques for estimation of the M_s-m_b line result in the smallest bias. Then these techniques should be applied to the real data, and any expected errors can be allowed for in the interpretation. Results from our particular network simulation indicate that maximum-likelihood estimation with symmetrical cutoffs on the data at the 90% LR threshold is an accurate technique and that regression of $\hat{m}_b$ on $\hat{M}_s$ with a 90% threshold cutoff on $\hat{M}_s$ only is a good alternative technique.

In an earlier theoretical section, a method of estimating the true M_s-m_b line was outlined; essentially it involves the use of equation (25). This method is valid for single-station plots only, and it also requires more precise knowledge of detection probabilities than maximum likelihood and standard regression methods do. An illustration of the method is worthwhile though. Using MSBNET simulated data for one station of the previously employed network, an incremental plot of number of events versus $\hat{m}_b$, shown in Figure 33, was made to determine the seismicity-magnitude relation and thus the expected number of events N_i at each m_{b_i}. (The low P threshold case with $\mu_b'=3.00$ was used.) The surface-wave and body-wave detection probabilities for this station, relative to station magnitude, are given by equation (2) with $\sigma_n=0.2$ and 50% probabilities at 3.80 and 3.00, respectively. In practice, these probabilities could be estimated from noise values or incremental detection plots. This information was applied in the evaluation of equation (25). A value of $\varepsilon=0.1$ was selected arbitrarily for one set of evaluations. Also a value of $\varepsilon=0.35$, as determined by equation (23), was used to attempt to estimate exactly the true M_s-m_b line. Figure 34 shows the MSBNET data and the calculated $(\hat{M}_{s_i},\hat{m}_{b_i})$ points. No evaluations were made for $\hat{m}_b<3.3$, which is roughly the .95 detection probability point for body waves as seen in Figure 33; and no summations were made which included a surface-wave detection probability less

than .16, at $\hat{M}_s$=3.6. The results appear to be excellent, with the ε=0.35 points closely defining the true M_s-m_b line. The line has been estimated down to m_b=3.3, the 95% body-wave threshold for the station. Note that the method should work reasonably well even if the true M_s-m_b relation were not linear. In that case the seismicity-magnitude relation would not be log-linear and the density function $f(\hat{M}_s|\hat{m}_b)$ given by equation (20) would no longer be strictly correct; but the deviations should not be large enough to nullify application of the procedure, illustrated here for a linear M_s-m_b relation, to somewhat non-linear relations.

A closely related M_s-m_b estimation procedure which does not require estimation of the seismicity-magnitude relation nor station M_s thresholds is possible for m_b values above the 90% detection threshold. In this case one simply sums down to an $\hat{M}_{s_i}$ which gives a fixed fraction ε of the events actually observed at $\hat{m}_{b_i}$. This assumes perfect detection down to $\hat{M}_{s_i}$. The resulting curve fit to the $(\hat{M}_{s_i},\hat{m}_{b_i})$ pairs may then be shifted down by the difference at high magnitudes between the $\hat{M}_{s_i}$'s obtained using ε and the $\hat{M}_{s_i}$'s obtained using an ε_m corresponding to the true M_s-m_b line ($\varepsilon_m \approx 0.5$ might not be too inaccurate for most practical purposes). We have evaluated this procedure using program MSBNET and find that it is apparently accurate, although somewhat unstable, for the case ε=0.1, which was chosen so that the estimation of the M_s-m_b line could be extended to low magnitudes.

Comparison of Simulation with LASA M_s-m_b Data for the Kamchatka-Kuril Region

Figure 35 shows the incremental and cumulative curves for the SAAC-LASA daily summary for geographic regions 217-222 in the Kamchatka-Kuril area for approximately six months in 1972. The α value is approximately 0.8 as in the simulation. The incremental and cumulative thresholds from this figure are given in Table VIII. Mack and Robertson (1973) gave a cumulative curve from which it is possible to obtain the cumulative 50% and 90% points for LASA LR detection in the Kamchatka-Kuril region for events having a P detection as reported in the LASA bulletins. They also showed a curve of direct probability of LR detection as a function of LASA $\hat{m}_b$.

In choosing our simulation parameters, we had these LASA data in mind; therefore we adjusted the mean single-station short-period and long-period noise levels so that the difference in simulated thresholds would be as observed at LASA, and we also chose the absolute levels of the noise values so that the simulated thresholds were in agreement with the observed values (to within .01 magnitude unit). Finally, setting $M_s=m_b-0.8$ instead of $M_s=m_b$ as in the simulation enables us to present in Table VIII a comparison between observed and simulated thresholds. In general we feel that the agreement is within the experimental error. More conclusive tests of the agreement of simulations with observation should be possible using

the LPL data. We shall treat this problem in future reports. It will be important to see if network incremental detection curves for small source regions do indeed have humps.

REFINEMENTS IN CALCULATED THRESHOLDS

S/N Ratio Required for Detection

A variation of equation (1)

$$P_D(m) = \Phi\left[\frac{\log A - \mu + \log r}{(\sigma_s^2 + \sigma_n^2)^{1/2}}\right] \qquad (28)$$

is most often used in predicting thresholds; here r is the S/N ratio required for detection. An assigned value of r must depend on the frequency band of the signal and of the noise, and on the analysis procedure. If events having nearly the same frequency content as the background noise are being searched for without knowledge of their time and frequency of occurrence, then a rather high r may be set for calculating detection thresholds. On the other hand, if events with frequency content different from the recorded noise and with accurate predicted arrival times are being search for, a rather low r must apply. It is mandatory that r be determined via a controlled experiment when calculated detection probabilities are to be compared with empirical probabilities determined in analysis of new seismic data. Such an experiment is reported by Geotech (1966) for short-period detection at the VELA observatories where a S/N ratio (peak-to-peak signal over peak-to-peak noise) of roughly 1.5 seemed appropriate for detection at the 50% level. Another experiment is reported by von Seggern (1974) for the Long-Period Experimental

network where a S/N ratio (peak-to-peak signal over peak-to-peak noise) of roughly 1.0 seem appropriate for LR detection at the 50% level. Note that the choice of definition of μ, peak-to-peak or rms, affects the value of r. There is approximately a factor of six difference in the two definitions, since if the noise is Gaussian, a peak-to-peak noise excursion six times greater than the rms is expected once for roughly every 7 minutes of long-period data sampled at 1 pt/sec, or seven times greater than the rms once for roughly every 33 minutes.

Background Noise Level

The value of μ in equation (28) should be determined empirically from actual noise recorded at the site. It is often taken as the rms value of a noise sample or average rms of many noise samples (e.g., Benno, 1972). For the purpose of detecting signals, however, there is only a tenuous connection between the rms value and the noise peaks which effectively set the detection threshold. The rule of thumb for multiplying rms by six to get peak-to-peak noise maxima only applies if the noise is Gaussian, this requirement being especially important for extreme values of the distribution. Therefore a simple visual assessment of peak noise values (Geotech, 1966; von Seggern, 1974) is at least as good as and in some cases undoubtedly better than rms measurements. This approach estimates exactly what maximum noise amplitudes the human or

machine detector is up against to reliably pick signals without creating an unacceptable false alarm rate.

Another matter to be considered in determining μ in equation (28) is the presence of interfering events on seismic recordings. Many events which probably would have been detected in normal background are masked by larger signals from another event. These interfering events should logically be considered as the upper values of noise. If μ is determined solely from quiescent periods, the resulting predictions of station or network capability using (28) will necessarily be lower than actually achieved. This bias is not insignificant. In Figure 36 we show two noise distributions for the Long-Period Experimental station at Charters Towers, Australia. Both distributions represent the picking of maximum peak-to-peak excursions of "noise" with period near 20 seconds on the vertical component in a 40-minute window at roughly equally-spaced dates in 1972; the samples with the lower mean were taken always at seismically inactive times, and the samples with the higher mean were taken at a fixed time of day no matter what was present on the recording. The logarithm of the means differ by nearly .3 magnitude unit. We suggest that the higher μ is the appropriate one to use in predicting thresholds if interfering events cannot be eliminated in detection studies.

SUMMARY AND CONCLUSIONS

In order to properly define seismic thresholds, it was necessary to define three levels of seismic magnitude: true, operational, and observed. The observed magnitudes are continually at our disposal, but we must determine thresholds on operational magnitude for the purpose of fair comparisons of detection capability among stations and networks. A further refinement into the frame of true magnitude ties thresholds to some absolute indicator of event magnitude, such as total energy release in explosions or earthquakes. Most of the difficulties in treating thresholds can be traced to the fact that natural seismicity results in more events scattering into a certain magnitude from lower magnitudes than from higher ones and to the fact that there is increasing magnitude bias with decreasing magnitude.

Many of our problems in threshold determination can be handled analytically. We established in this manner that:

1. For a single station, the 50% incremental threshold magnitude using observed magnitudes is the same as the 50% direct threshold determination using operational magnitudes; but the 90% and 10% are lower and higher, respectively.

2. The single-station incremental curve on observed $\hat{m}$ smoothly dies off, but the network curve has a hump near the threshold.

3. The seismicity-magnitude relation determined using observed magnitude will always be displaced upward in magnitude compared to the true seismicity.

4. Cumulative threshold magnitudes, either network or single-station, will always be lower than the incremental ones.

5. The distribution of magnitudes along a straight line (horizontal or vertical) through a population of M_s-m_b data is dependent on the seismicity-magnitude relation and, although normal, has varying parameters.

We have presented the method for converting seismic thresholds for LR from M_s to m_b scales and vice versa. Because it is necessary to invoke several estimated parameters in this procedure, such as the amount of source bias, the extent of path effects on amplitude, the slope and intercept of the M_s-m_b relation, and the slope and intercept of the seismicity-magnitude relation, we feel this conversion should be done with care, lest results stray from predictions.

We have assessed the effect of source bias and noise correlation on detection thresholds using a modified version (NETWCORR) of our standard program for predicting network detection capability. Results show that network capabilities are altered insignificantly by noise correlation and that predicted thresholds for true and operational magnitudes typically differ by roughly .1 magnitude unit.

An accurate method for reconstructing the true M_s-m_b line from single station data was derived theoretically

using the statistical properties of the observed data compared to the total population of possible data.

A simulation computer program was written to gain insights in such matters as network threshold problems not analytically tractable, effects of correlated signals and noise, feedback between P and LR thresholds, and estimation of M_s-m_b relationships. Through the analysis of an ideal network we have attempted to illustrate the range of situations which arise in threshold determination. Comparison of program MSBNET results with those of real networks enables one to infer certain parameters for the network and to make threshold corrections to the empirical data.

Simulation results agreed with theoretical predictions wherever comparison was possible. The simulation experiment established these important facts:

1. Varying the P threshold does not affect the single station direct LR threshold on $\hat{m}_b$ but does affect the network threshold. Incremental and cumulative thresholds can be strongly affected in either case.

2. The hump in the network LR incremental detection curve is reduced by assuming significant source bias and also by plotting versus $\hat{m}_b$ rather than $\hat{M}_s$.

3. The 90% single-station LR thresholds in terms of $\hat{m}_b$ are typically much greater than the thresholds in terms of $\hat{M}_s$ (assuming $M_s = m_b$) of the order of 0.4 magnitude units. In general, really substantial effects are possible and care must be exercised when comparing thresholds determined in different ways.

4. Noise correlated between stations at realistic levels (ρ=0.4) has an almost negligible effect on thresholds in the presence of realistic levels of source bias.

5. The seismicity bias for large magnitudes is negligible if network magnitudes are used, except in the presence of source bias in which case it may be on the order of 0.1 magnitude unit.

6. For realistic network parameters, compared to the 1/10 thresholds, the 2/10 thresholds increase by 0.10-0.15 magnitude unit and, compared to the 2/10 thresholds, the 4/10 thresholds increase by a similar amount. This is in rough agreement with, but slightly less than, the NETWORTH differentials of 0.15-0.20.

7. Use of the error function as a probability of detection curve for a network can in some cases be a poor approximation to the curve deduced by simulation.

8. The true M_s-m_b relationship is accurately estimated by the maximum likelihood method if a suitable low cutoff is made on the M_s and m_b values. Regression of m_b on M_s also yields accurate results if that 10% of the events with the lowest M_s values are excluded from the calculation.

As an example of the usefulness of a simulation approach to threshold problems, good agreement was achieved between simulated results from MSBNET and LASA data for LR and P detection.

In regard to predicting seismic thresholds, we reiterate the importance of determining realistic background "noise" values which take into account seismic waves from known sources and also the importance of determining the S/N ratio required for detection in a specified context.

ACKNOWLEDGEMENTS

The programming efforts of several individuals helped to produce the results of this report. H. L. Husted modified M. Wirth's program NETWORTH to account for signal and noise correlation; H. L. Husted and T. W. McElfresh developed the simulation program MSBNET; and D. D. Nelson programmed some of the theoretical calculations. Discussions with R. H. Shumway on some of the probability theory in this report are also gratefully acknowledged.

REFERENCES

Abramowitz, M. and Stegun, I. A., 1964, Handbook of Mathematical Functions, National Bureau of Standard, United States Department of Commerce.

Benno, S. A., 197?, Preliminary evaluation of single stations of the Long Period Experimental Network, Special Report No. 7 Extended Array Evaluation Program, Texas Instruments, Inc., Dallas, Texas.

Blandford, R. R. and Wirth, M. H., 1973, Automatic array and network detection in the presence of signal variability, Seismic Data Laboratory Report No. 308, Teledyne Geotech, Alexandria, Virginia.

Chang, A. C., 1973, A comparison of the LASA-NORSAR short period arrays, SAAC Report No. 12, Teledyne Geotech, Alexandria, Virginia.

Ellison, B. E., 1964, Two theorems for inferences about the normal distribution with application in acceptance sampling, J. Amer. Statist. Assoc., v. 59, p. 84-96.

Ericsson, U., 1971a, A linear model for the yield dependence of magnitudes measured by a seismographic network, Geophys. J. R. Astr. Soc., v. 25, p. 49-70.

Ericsson, U., 1971b, Maximum likelihood linear fitting when both variables have normal and correlated errors, FOA-4 Report NO. C-4474-A1, Research Institute of National Defense, Stockholm, Sweden.

REFERENCES (Continued)

Freedman, H., 1967, Estimating earthquake magnitude, Bull. Seismol. Soc. Amer., v. 57, p. 747-760.

Geotech, 1966, Estimates of the detection capability of four VELA-Uniform seismological observatories, Technical Report No. 66-1, Teledyne Geotech, Garland, Texas.

Gutenberg, B. and Richter, C. F., 1954, Seismicity of the Earth, 2nd edition, Princeton University Press, Princeton, New Jersey.

Harley, T. W. and Heiting, L. N., 1972, Indirect estimates of surface wave detection probabilities, Special Report No. 1, Extended Array Evaluation Program, Texas Instruments, Inc., Dallas, Texas.

Herrin, E. and Tucker, W., 1972, On the estimation of body-wave magnitudes, Report to the Air Force Office of Scientific Research, Dallas Geophysical Laboratory, Southern Methodist University, Dallas, Texas.

Kelley, E. J. and Lacoss, R. T., 1969, Estimation of seismicity and network detection capability, Technical Note 1969-41, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts.

Lacoss, R. T., 1972, Seismic event detection and discrimination-some statistical considerations, in

Proceedings from the Seminar on Seismology and Seismic Arrays, edited by E. S. Husebye and H. Bunyum, NTNF/NORSAR, Kjeller, Norway.

Mack, H., 1971, Evaluation of the large array long-period network, SAAC Report No. 4, Teledyne Geotech, Alexandria, Virginia.

Mack, H. and Robertson, H. C., 1973, Detection threshold of the LASA, ALPS, NORSAR long-period network, SAAC Report No. 10, Teledyne Geotech, Alexandria, Virginia.

Scholtz, C. H., 1968, The frequency-magnitude relation of microfracturing in rock and it relation to earthquakes, Bull. Seismol. Soc. Amer., v. 58, p. 399-416.

Shumway, R. H. and Blandford, R. R., 1970, A simulation of seismic discriminant analysis, SDL Report No. 261, Teledyne Geotech, Alexandria, Virginia.

von Seggern, D. H., 1972, Joint magnitude determination and the analysis of variance for explosion magnitude estimates, SDL Report No. 286, Teledyne Geotech, Alexandria, Virginia.

von Seggern, D. H., 1974, Final report on the analysis of recordings from the Very Long Period Experimental stations, Teledyne Geotech, Alexandria, Virginia.

REFERENCES (Continued)

Wirth, M. H., 1970, Estimation of network detection and location capability, SDL Research Memorandum, Teledyne Geotech, Alexandria, Virginia.

Zachs, S. and Even, M., 1966, The efficiencies in small samples of the maximum likelihood and best unbiased estimators of reliability functions, J. Amer. Statist. Assoc., v. 61, p. 1033-1051.

TABLE I

Bias of Single-Station Incremental Threshold
Relative to Operational NETWORTH Threshold

		σ_s											
		.1			.2			.3			.4		
		σ_n			σ_n			σ_n			σ_n		
		.1	.2	.3	.1	.2	.3	.1	.2	.3	.1	.2	.3
Threshold	10%	.05	.03	.02	.16	.11	.08	.28	.20	.16	.40	.29	.25
	50%	0	0	0	0	0	0	0	0	0	0	0	0
	90%	-.05	-.03	-.02	-.16	-.11	-.08	-.28	-.20	-.16	-.40	-.29	-.25

TABLE II

Bias of Single-Station Cumulative Threshold
Relative to Incremental Threshold

		σ_n of $P_I(\hat{m})$											
		.1			.2			.3			.4		
		α			α			α			α		
		1.0	1.5	2.0	1.0	1.5	2.0	1.0	1.5	2.0	1.0	1.5	2.0
Threshold	10%	-.84	-.52	-.37	-.74	-.44	-.32	-.70	-.42	-.30	-.63	-.38	-.26
	50%	-.26	-.19	-.15	-.31	-.22	-.17	-.34	-.23	-.18	-.35	-.24	-.18
	90%	-.11	-.10	-.08	-.18	-.14	-.11	-.21	-.16	-.13	-.25	-.17	-.15

TABLE III

Simulation Parameters for Network of 10 Equal Capability Stations

$\sigma_{mb} = 0.212$ $\quad$ $\sigma_{sb} = 0.212$

$\sigma_{ms} = 0.212$ $\quad$ $\sigma_{ss} = 0.212$

$\sigma_n = 0.2$ $\quad$ $\rho_n = 0.4$ (when ρ_n assumed $\neq 0$)

$\mu_b' = 3.47$

$\mu_s' = 3.80$ (assumes $M_s = m_b$)

$\alpha = 0.8$

$a = 1.0$

$b = 0.0$

Note: $\sqrt{\sigma_{mb}^2 + \sigma_{sb}^2} = 0.3$ and $\sqrt{\sigma_{ms}^2 + \sigma_{ss}^2} = 0.3$.

Note: if σ_{mb} and/or $\sigma_{ms} = 0$; then σ_{sb} and/or $\sigma_{ss} = 0.3$.

TABLE IV

Predicted Thresholds for the Simulated Network
Looking at Kuril-Kamchat a Region

System	Source Bias	Correlated Noise	Magnitude Reference	10%	50%	90%
Single Station	no	no	T, OP	3.34	3.80	4.26
Single Station	yes	no	T	3.34	3.80	4.26
Single Station	yes	no	OP	3.44	3.80	4.16
1/10 (Standard Networth)	no	no	T, OP	2.95	3.26	3.51
	yes	no	T	2.99	3.36	3.70
	yes	no	OP	3.13	3.37	3.57
	yes	yes	T	3.04	3.41	3.74
	yes	yes	OP	3.14	3.40	3.64
2/10 (Standard Networth)	no	no	T, OP	3.22	3.45	3.65
	yes	no	T	3.21	3.50	3.83
	yes	no	OP	3.34	3.51	3.68
	yes	yes	T	3.19	3.54	3.87
	yes	yes	OP	3.30	3.54	3.76
4/10 (Standard Networth)	no	no	T, OP	3.48	3.67	3.85
	yes	no	T	3.38	3.68	4.02
	yes	no	OP	3.55	3.69	3.84
	yes	yes	T	3.36	3.69	4.04
	yes	yes	OP	3.47	3.70	3.91

Note: Magnitudes are M_s - for m_b subtract 0.33

Note: Source bias creates a difference between true and operational magnitude even when all stations detect easily.

Note: See Table III for Network parameters.

TABLE V

Observed and Predicted Thresholds with Source Bias

	Threshold in terms of:	Figure No:	Incremental		Cumulative		Direct		True Magnitude, Standard NETWORTH	
			50%	90%	50%	90%	50%	90%	50%	90%
Single Station Thresholds										
1. NETWORTH Single Station Det. of P	$\tilde{m}_b$	7	x	x	x	x	x	x	3.47	3.93
2. NETWORTH Single Station Det. of LR	$\tilde{M}_s$	7	x	x	x	x	x	x	3.80	4.26
3. Single Station Det. of P	$\hat{m}_b$	8	3.47	3.71	3.12	3.55	x	x	x	x
4. Single Station Det. of LR	$\hat{M}_s$	8	3.80	4.04	3.45	3.88	x	x	x	x
5. Single Station Det. of LR Low Single Station P Det. Threshold (3.00)	$\hat{m}_b$	12,11	3.94	4.44	3.54	4.25	3.95	4.57	x	x
6. Single Station Det. of LR|Standard Single Station P Det. Threshold (3.47)	$\hat{m}_b$	10,11	3.94	4.44	3.57	4.25	3.97	4.56	x	x
7. Single Station Det. of LR|Standard Single Station P Det. Threshold (3.47)	$\hat{M}_s$	9	3.83	4.06	3.45	3.89	x	x	x	x
Network Thresholds										
8. NETWORTH 1/10 Det. of LR	$\tilde{M}_s$	7	x	x	x	x	x	x	3.36	3.70
9. NETWORTH 2/10 Det. of LR	$\tilde{M}_s$	7	x	x	x	x	x	x	3.50	3.83
10. NETWORTH 4/10 Det. of LR	$\tilde{M}_s$	7	x	x	x	x	x	x	3.68	4.02
11. NETWORTH 4/10 Det. of P	$\tilde{m}_b$	7	x	x	x	x	x	x	3.35	3.69
12. 4/10 Detection of P	$\hat{m}_b$	13	3.46	3.51	3.00	3.33	x	x	x	x
13. 1/10 Detection of LR|Low P Threshold (3.00), 4/10 Det. of P	$\hat{m}_b$	14,11	3.37	3.75	3.00	3.48	3.49	3.90	x	x
14. 1/10 Detection of LR|Low P Threshold (3.00), 4/10 Det. of P	$\hat{M}_s$	15	3.57	3.65	3.05	3.37	x	x	x	x
15. 1/10 Detection of LR|4/10 Det. of P	$\hat{m}_b$	16a,11	3.55	3.71	3.21	3.56	3.62	3.97	x	x
16. 1/10 Detection of LR|4/10 Det. of P	$\hat{M}_s$	17a	3.66	3.75	3.19	3.51	x	x	x	x
17. 2/10 Detection of LR|4/10 Det. of P	$\hat{m}_b$	18,11	3.63	3.84	3.30	3.65	3.75	4.08	x	x
18. 2/10 Detection of LR|4/10 Det. of P	$\hat{M}_s$	19	3.75	3.85	3.33	3.65	x	x	x	x
19. 4/10 Detection of LR|4/10 Det. of P	$\hat{m}_b$	20,11	3.78	4.05	3.45	3.87	3.86	4.20	x	x
20. 4/10 Detection of LR|4/10 Det. of P	$\hat{M}_s$	21	3.85	3.94	3.45	3.77	x	x	x	x
21. 1/10 Detection of LR|4/10 Det. of P $\rho_n = .40$	$\hat{m}_b$	16b,11	3.56	3.71	3.25	3.57	3.69	4.02	x	x
22. 1/10 Detection of LR|4/10 Det. of P $\rho_n = .40$	$\hat{M}_s$	17b	3.68	3.79	3.26	3.58	x	x	x	x

Note: See Table III for network parameters.

TABLE VI

Observed and Predicted Thresholds with No Source Bias

Thresholds in terms of:		Figure No.	Incremental		Cumulative		Direct		True Magnitude, Standard NETWORTH	
			50%	90%	50%	90%	50%	90%	50%	90%
Single Station Thresholds										
1. NETWORTH Single Station Det. of P	$\tilde{m}_b$	7	+	+	+	+	+	+	3.47	3.93
2. NETWORTH Single Station Det. of LR	$\tilde{M}_s$	7	+	+	+	+	+	+	3.80	4.26
3. Single Station Det. of P	$\hat{m}_b$	8	3.47	3.71	3.12	3.55	+	+	+	+
4. Single Station Det. of LR	$\hat{M}_s$	8	3.80	4.04	3.45	3.88	+	+	+	+
5. Single Station Det. of LR|Low Single Station P Det. Threshold (3.00)	$\hat{m}_b$	12,11	3.94	4.44	3.54	4.25	3.97	4.56	+	+
6. Single Station Det. of LR|Standard Single Station P Det. Threshold (3.47)	$\hat{m}_b$	10,11	3.94	4.44	3.57	4.25	3.97	4.56	+	+
7. Single Station Det. of LR|Standard Single Station P Det. Threshold (3.47) (3.47)	$\hat{M}_s$	9	3.83	4.06	3.45	3.89	+	+	+	+
Network Thresholds										
8. NETWORTH 1/10 Det. of LR	$\tilde{M}_s$	7	+	+	+	+	+	+	3.26	3.51
9. NETWORTH 2/10 Det. of LR	$\tilde{M}_s$	7	+	+	+	+	+	+	3.45	3.65
10. NETWORTH 4/10 Det. of LR	$\tilde{M}_s$	7	+	+	+	+	+	+	3.67	3.85
11. NETWORTH 4/10 Det. of P	$\tilde{m}_b$	7	+	+	+	+	+	+	3.34	3.52
12. 4/10 Detection of P	$\hat{m}_b$	23	3.49	3.56	2.98	3.30	+	+	+	+
13. 1/10 Detection of LR|Low P Threshold (3.00), 4/10 Det. of P	$\hat{m}_b$	24,22	3.31	3.46	2.98	3.31	3.44	3.63	+	+
14. 1/10 Detection of LR|Low P Threshold (3.00), 4/10 Det. of P	$\hat{M}_s$	25	3.57	3.65	2.97	3.27	+	+	+	
15. 1/10 Detection of LR|4/10 Det. of P	$\hat{m}_b$	26,22	3.57	3.65	3.10	3.42	3.56	3.80	+	+
16. 1/10 Detection of LR|4/10 Det. of P	$\hat{M}_s$	27	3.70	3.78	3.15	3.45	+	+	+	+
17. 2/10 Detection of LR|4/10 Det. of P	$\hat{m}_b$	28,22	3.66	3.74	3.25	3.57	3.72	3.92	+	+
18. 2/10 Detection of LR|4/10 Det. of P	$\hat{M}_s$	29	3.75	3.82	3.21	3.53	+	+	+	+
19. 4/10 Detection of LR|4/10 Det. of P	$\hat{m}_b$	30,22	3.80	3.90	3.44	3.76	3.88	4.12	+	+
20. 4/10 Detection of LR|4/10 Det. of P	$\hat{M}_s$	31	4.00	4.07	3.53	3.85	+	+	+	+

Note: See Table III for network parameters.

TABLE VII

Various Estimates of the Parameters a and b in the Linear Relation
$M_s = am_b + b$ Using the Network Simulation Data

		M_s Range															
		I		II		III		IV*		V*		VI*		VII**		VIII**	
		3.0 - 6.0		4.0 - 6.0		3.0 - 5.0		T_{90} - 6.0		T_{80} - 6.0		P_{90} - 6.0		P_{90} - 6.0		T_{90} - 6.0	
		a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b
1/10	Regression on $\hat{m}_b$	.79	.91	.81	.86	.67	1.39	.82	.77	.82	.78	.79	.90	.68	1.47	.74	1.18
	Regression on $\hat{M}_s$	1.08	- .27	1.23	-1.05	1.28	-1.00	1.16	- .63	1.19	- .78	1.10	- .36	1.00	.10	1.03	- .04
	Maximum Likelihood	.91	.44	1.00	.01	.89	.50	.96	.15	.98	.09	.91	.31	.79	.98	.85	.73
2/10	Regression on $\hat{m}_b$	.80	.91	.81	81	.61	1.65	.81	.82	.82	.78	.81	.85	.67	1.55	.75	1.10
	Regression on $\hat{M}_s$	1.08	- .26	1.23	-1.04	1.25	- .95	1.14	- .54	1.18	- .75	1.12	- .47	1.01	.04	1.06	- .18
	Maximum Likelihood	.91	.42	1.00	.01	.83	.78	.95	.23	.98	.10	.94	.28	.79	1.04	.87	.61
4/10	Regression on $\hat{m}_b$	.75	1.12	.81	.81	.56	1.93	.81	.84	.81	.81	.80	.87	.65	1.69	.70	1.34
	Regression on $\hat{M}_s$	1.04	- .12	1.23	-1.04	1.25	- .95	1.22	- .95	1.23	-1.04	1.16	- .69	1.04	- .12	1.03	- .16
	Maximum Likelihood	.90	+ .64	1.00	.02	.76	1.07	.99	.05	1.00	.02	.96	.20	.78	1.08	.83	.78

T_{90}, T_{80} - that M_s above which lie 90% or 80% of total events detected by LR.
P_{90} - that M_s which corresponds to the m_b at which 90% detection of LR occurs.
* Symmetrical cutoffs - m_b range is identical to M_s range.
** Nonsymmetrical cutoffs - m_b range is 3.0-6.0 throughout.

TABLE VIII

Predicted and Observed LASA Kuril-Kamchatka Thresholds

		Incremental		Cumulative		Direct on $\hat{m}_b$ **	
		50%	90%	50%	90%	50%	90%
m_b	Theoretical	3.46*	3.70	3.11	3.54	+	+
	Observed	3.46	3.66	3.15	3.48	+	+
M_s	Theoretical	3.00	3.25	2.64*	3.08	3.96	4.55
	Observed	NA	NA	2.64	3.00	3.95	4.40

*Forced fits to observation in effect determining mean noise levels.

**Assumes a P detection as reported in LASA bulletins.

$M_s = m_b - 0.8$ is assumed for conversion from simulation where $M_s = m_b$.

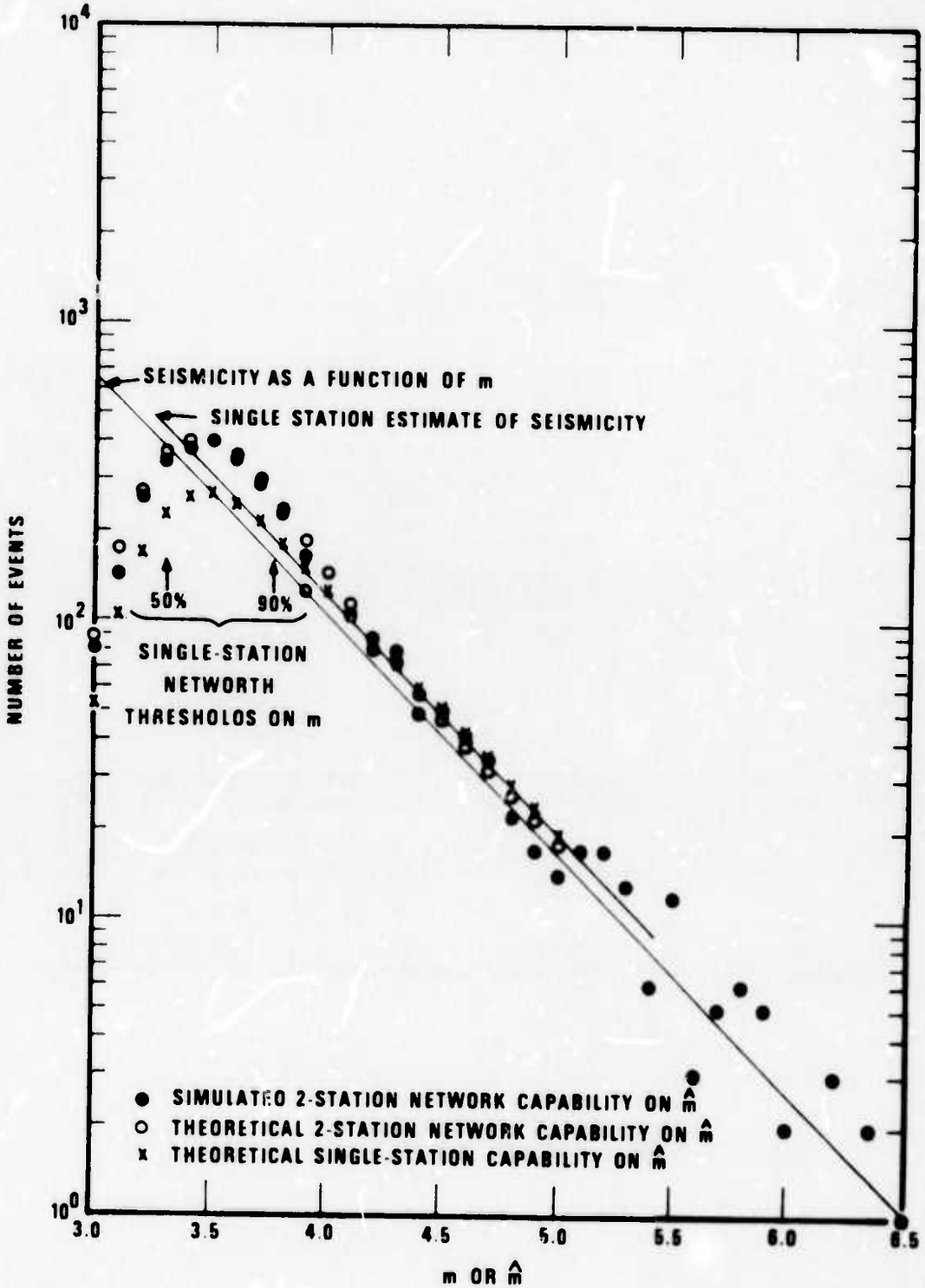


Figure 1. Incremental frequency-magnitude plots for a hypothetical data set with true, station-observed, and network-observed magnitudes as the basis (no source bias).

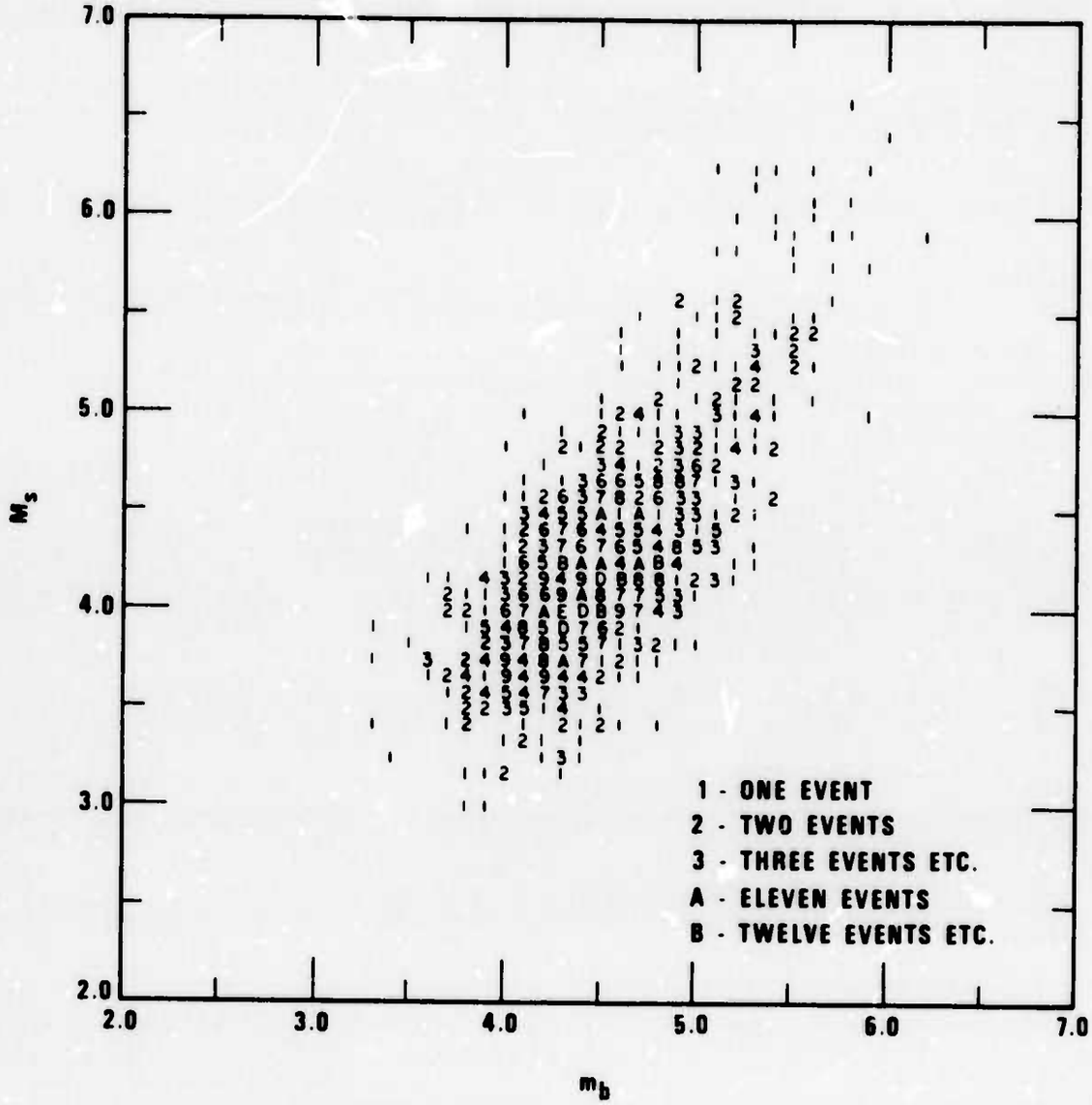


Figure 2. Typical M_s-m_b plot for earthquakes.

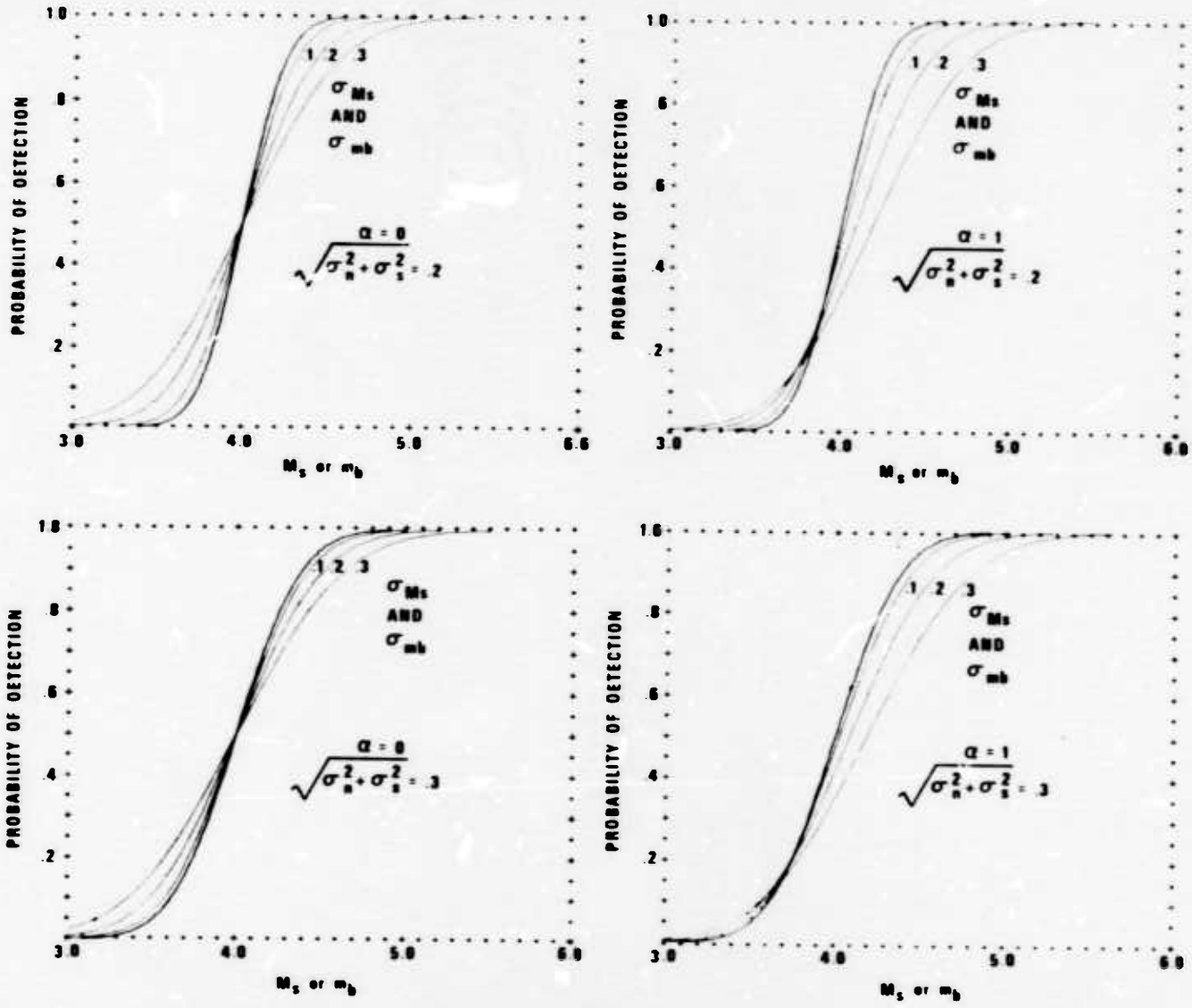


Figure 3. Predicted, direct, probability of detection curves on m_b derived from predicted, direct, probability of detection curves on M_s.

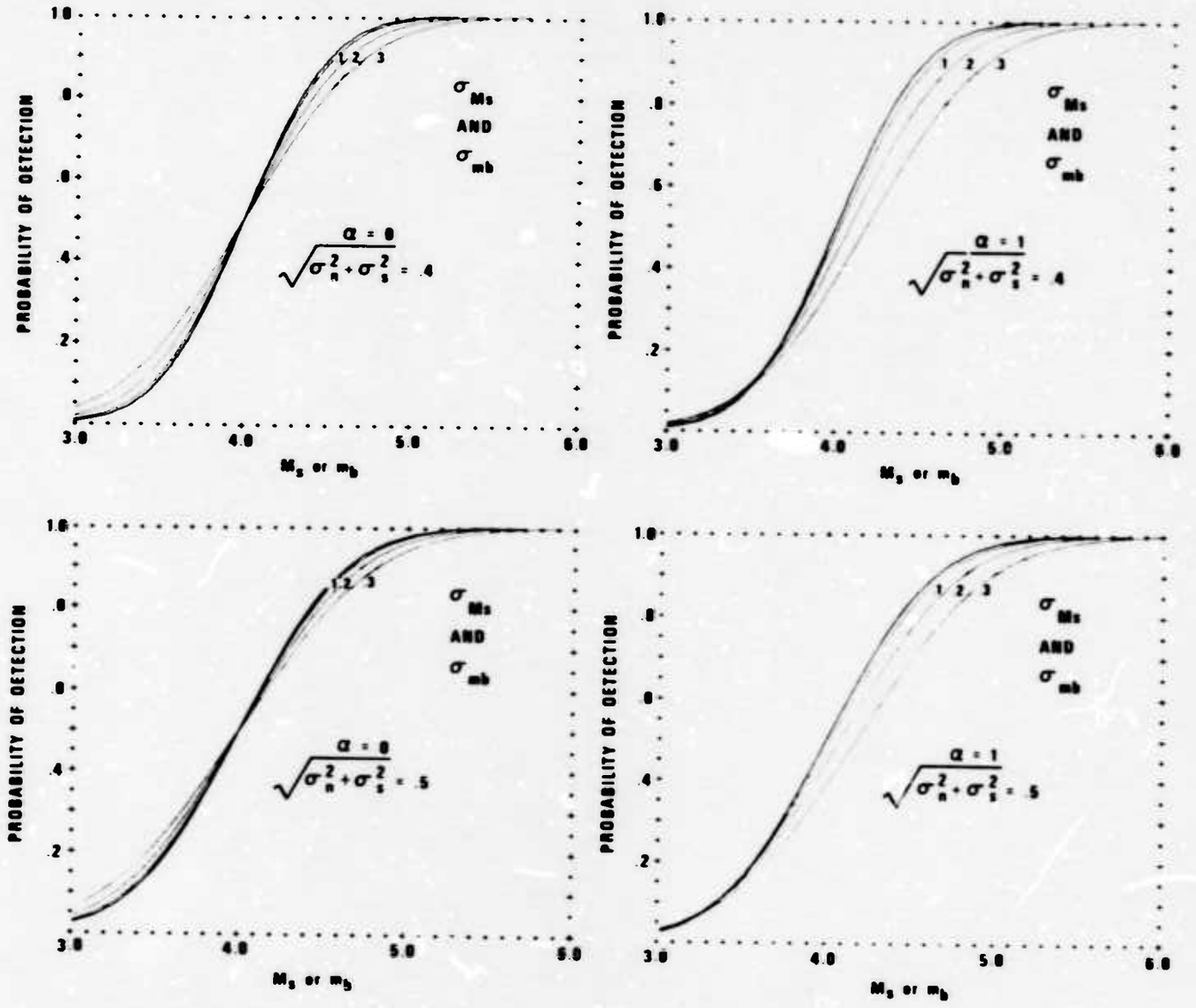


Figure 3 (Cont.). Predicted, direct, probability of detection curves on m_b derived from predicted, direct, probability of detection curves on M_s.

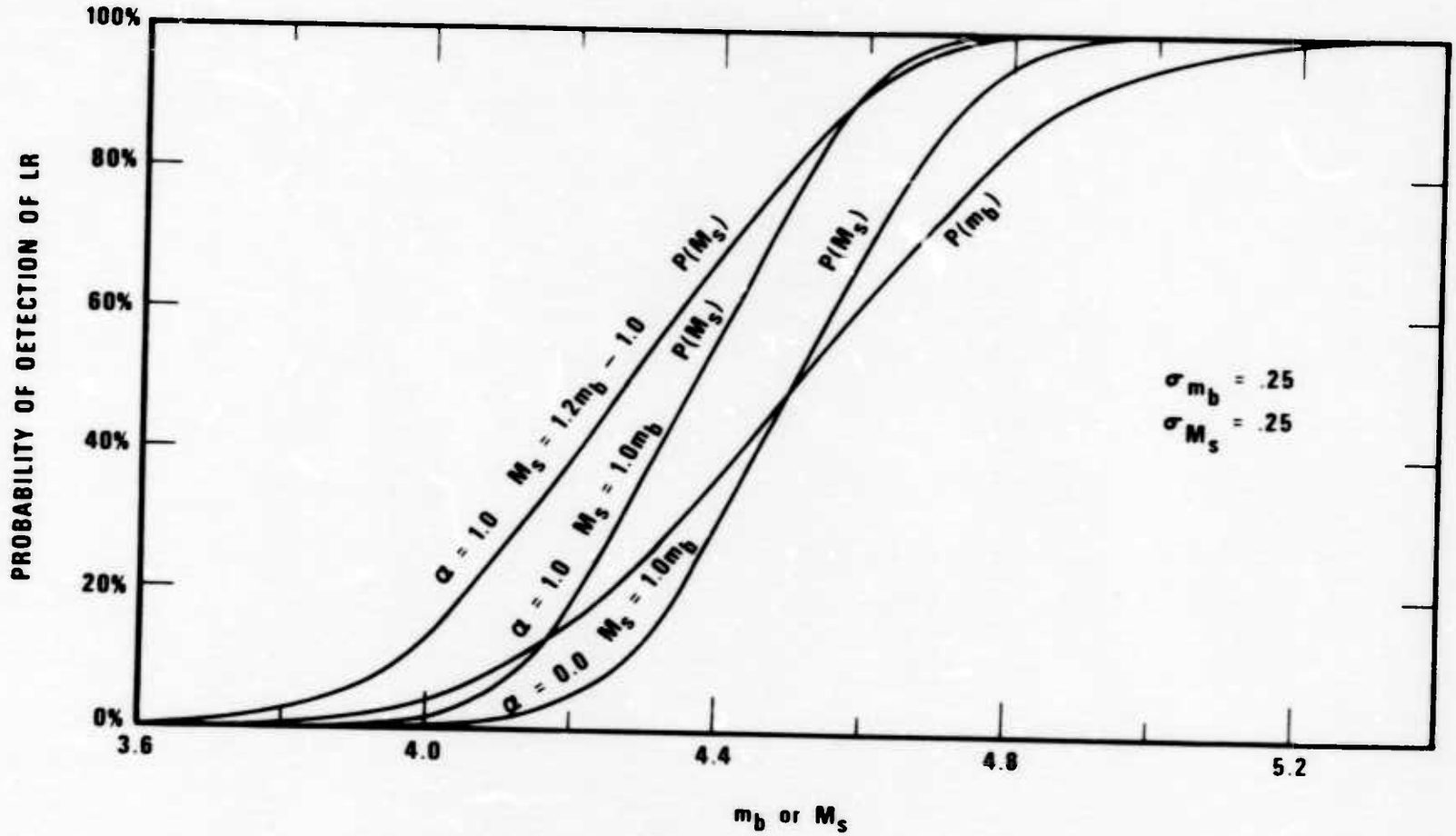


Figure 4. Transformation of surface-wave probability of detection curve on m_b to surface-wave probability of detection curve on M_s.

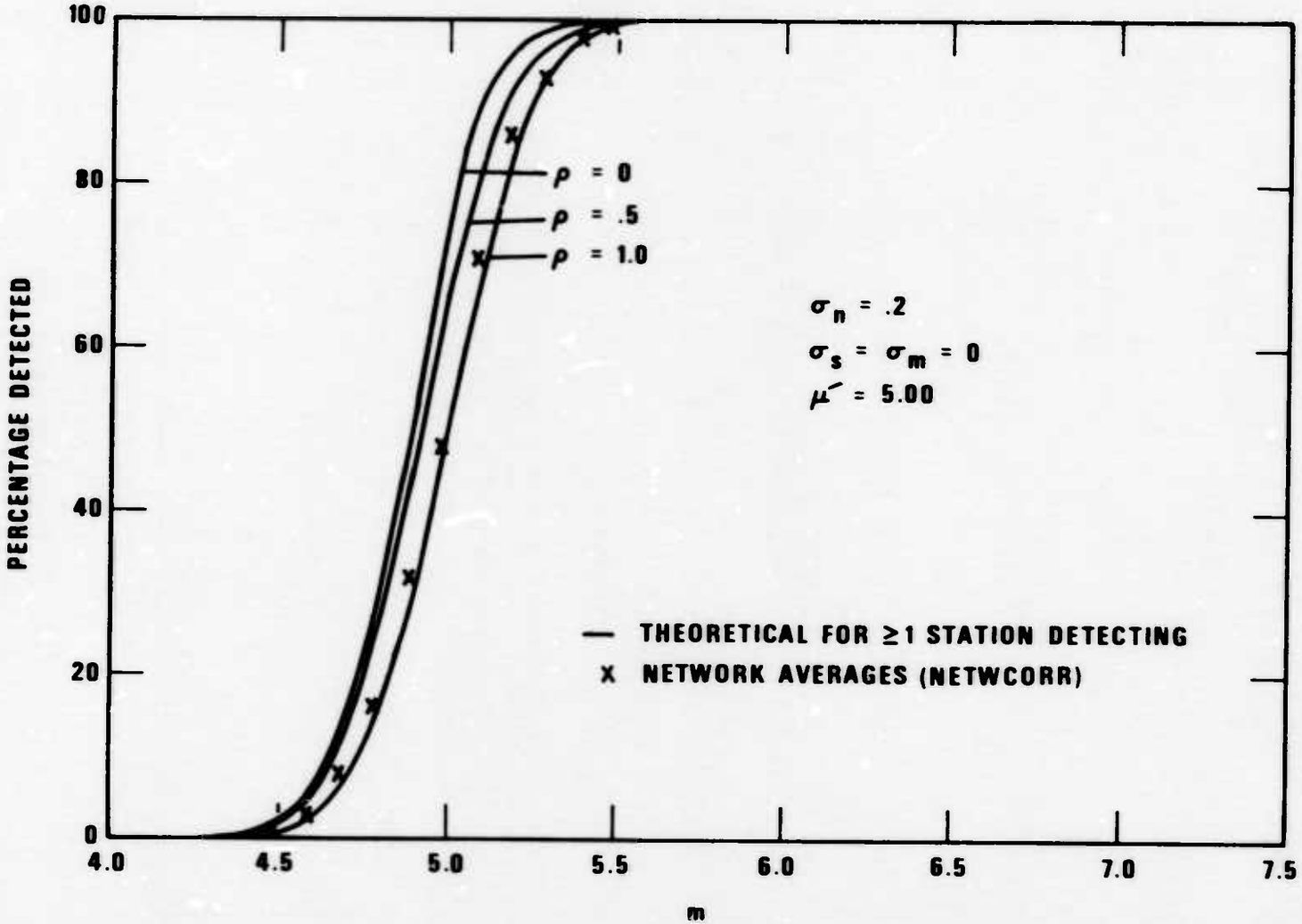


Figure 5. Theoretical effect of correlated signals and noise on a two-station network.

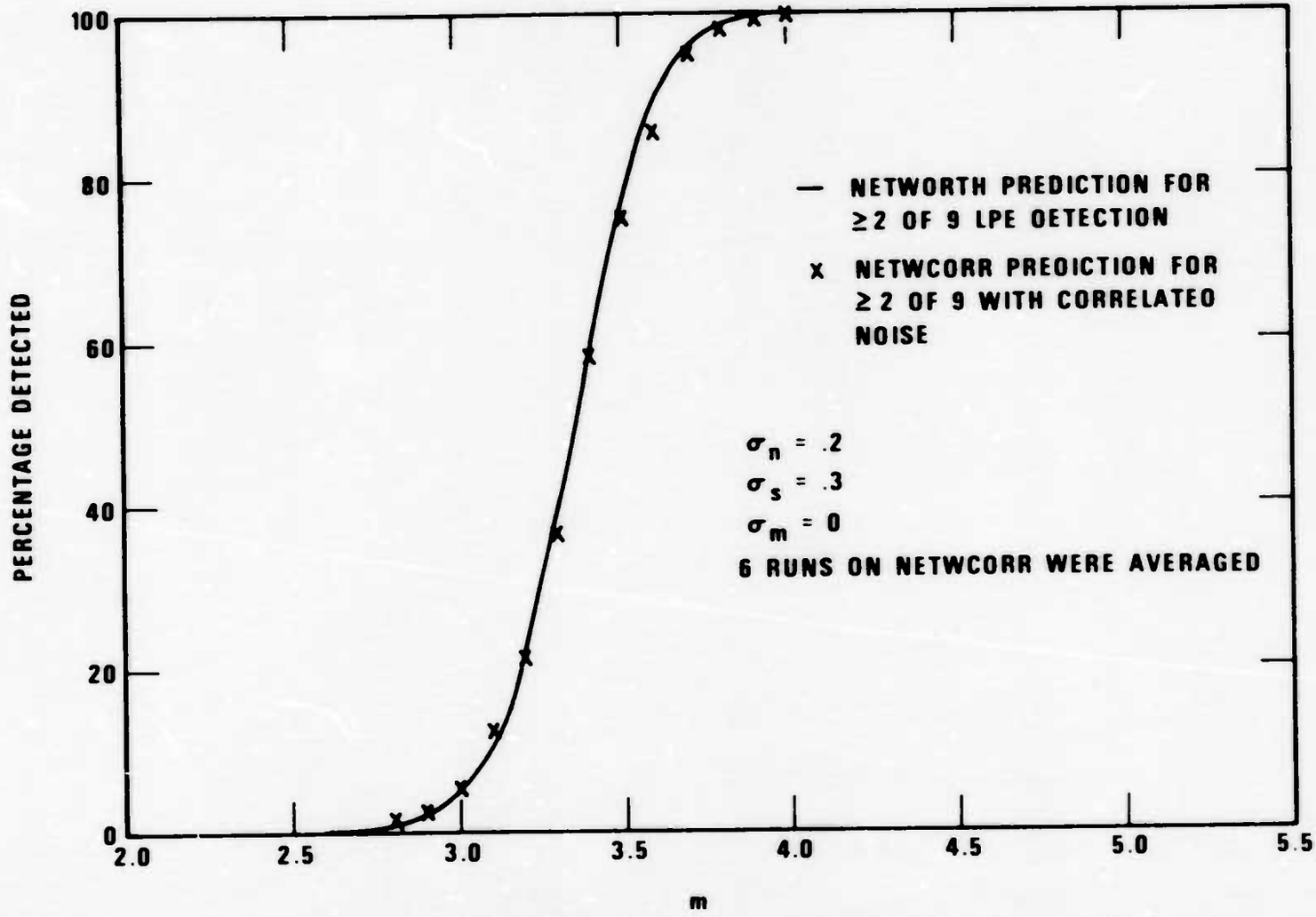


Figure 6. Predicted effect of correlated noise on the Long Period Experimental network.

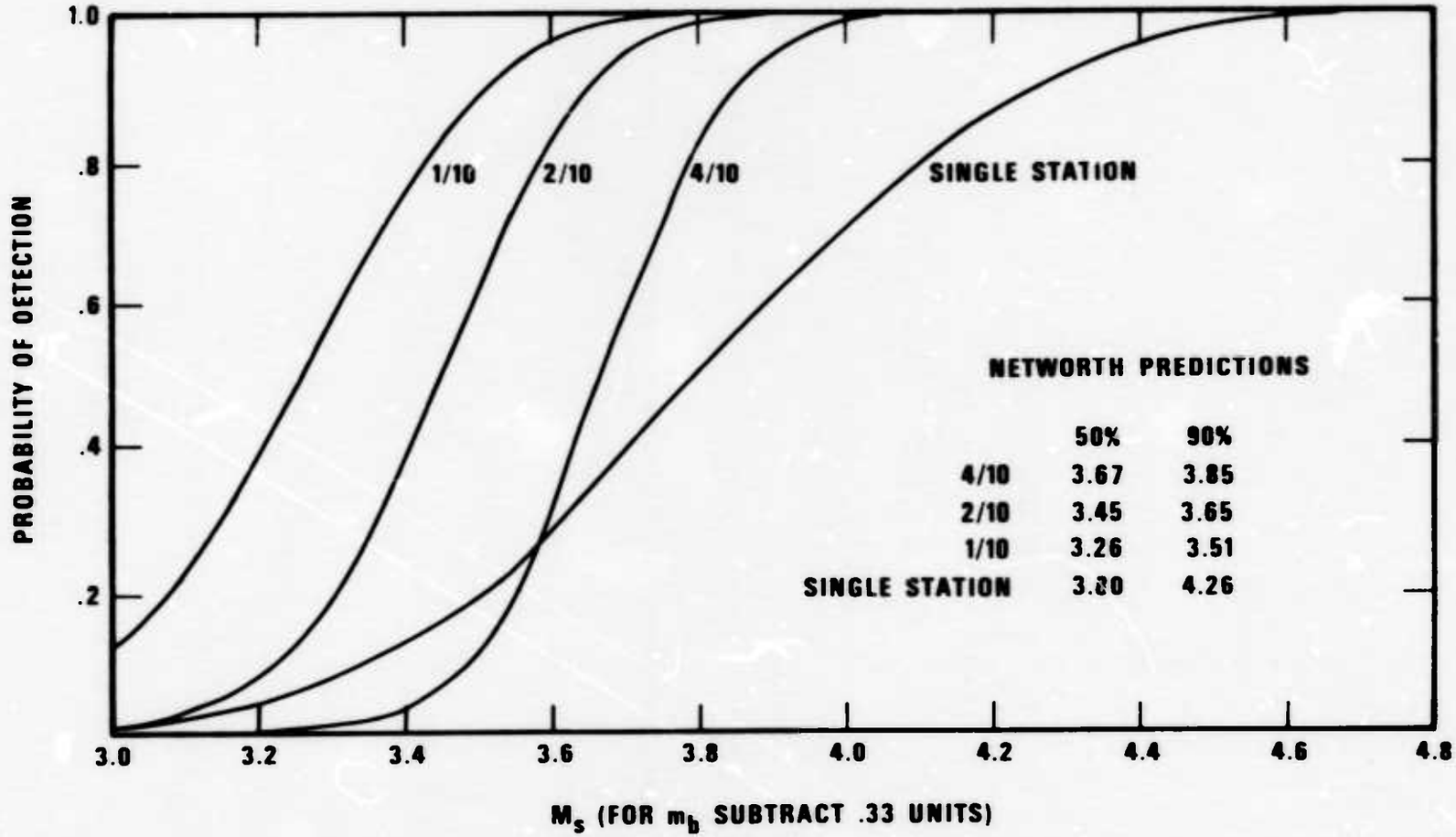


Figure 7. NETWORTH capability estimates for 10 LASA-type stations equally spaced about Kurii-Kamchatka with no source bias present.

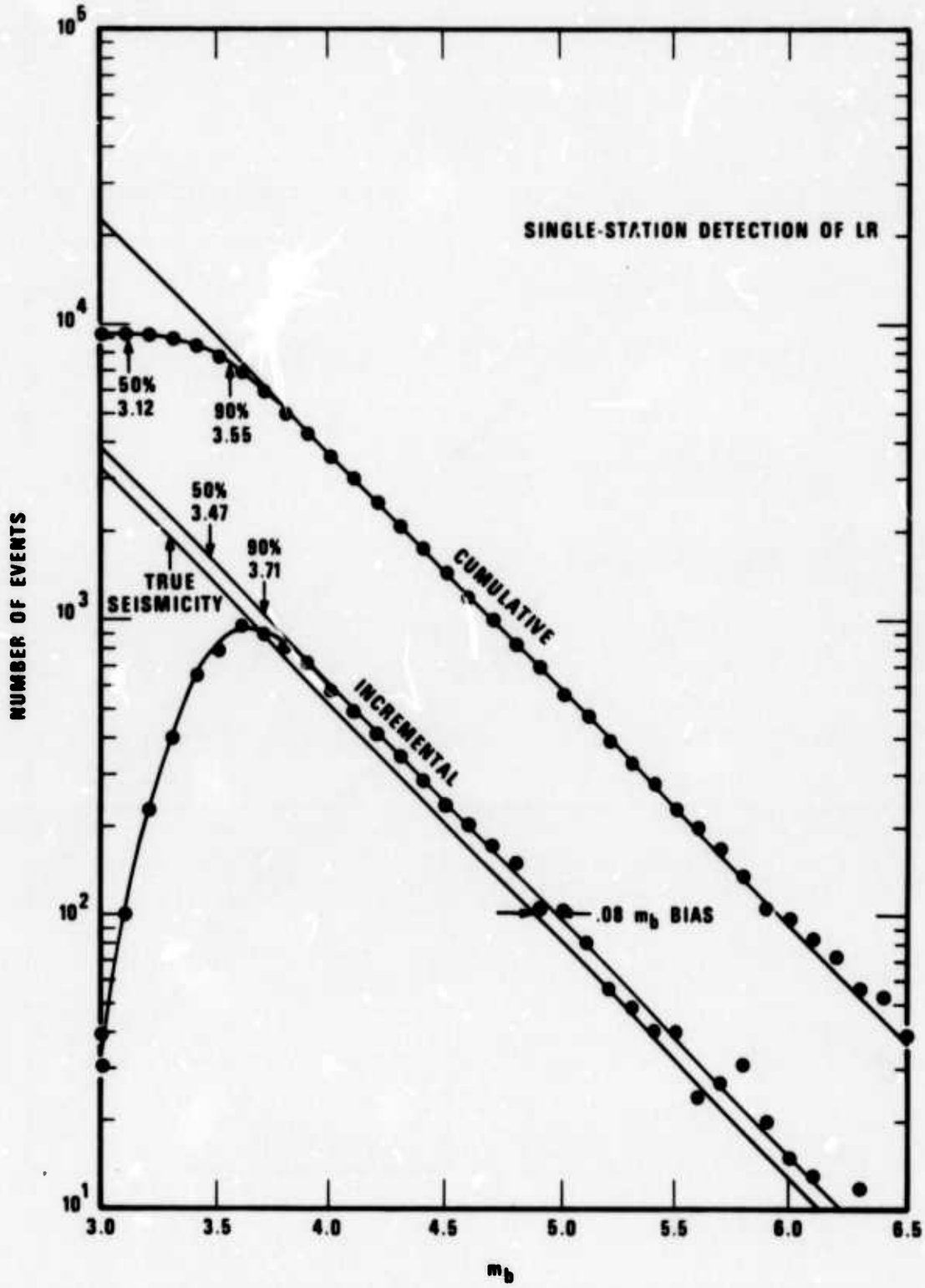


Figure 8. Simulated observed thresholds for LASA-type array looking at Kuril-Kamchatka region.

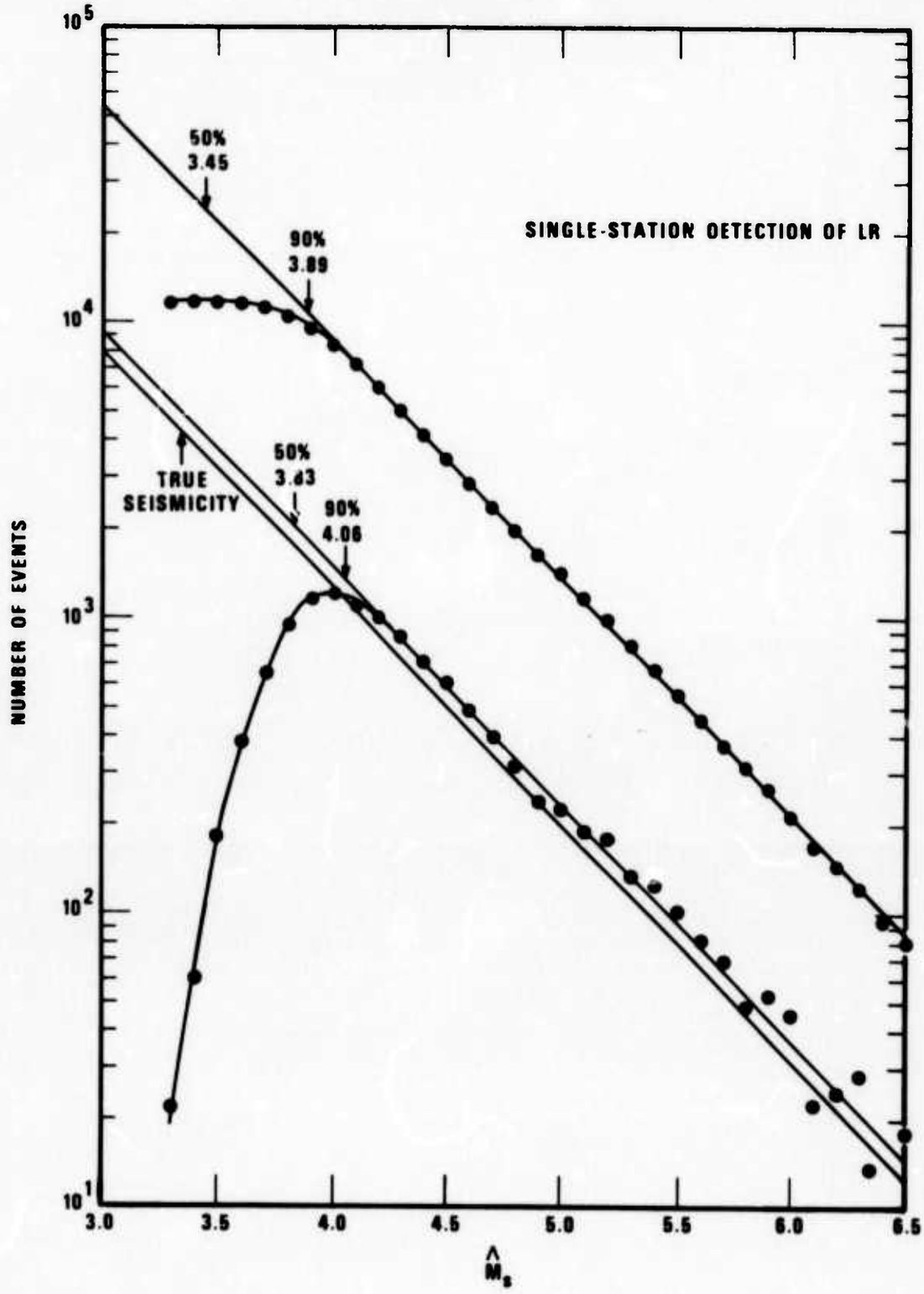


Figure 9. Single-station LR thresholds on $\hat{M}_s$ with P threshold of 3.47.

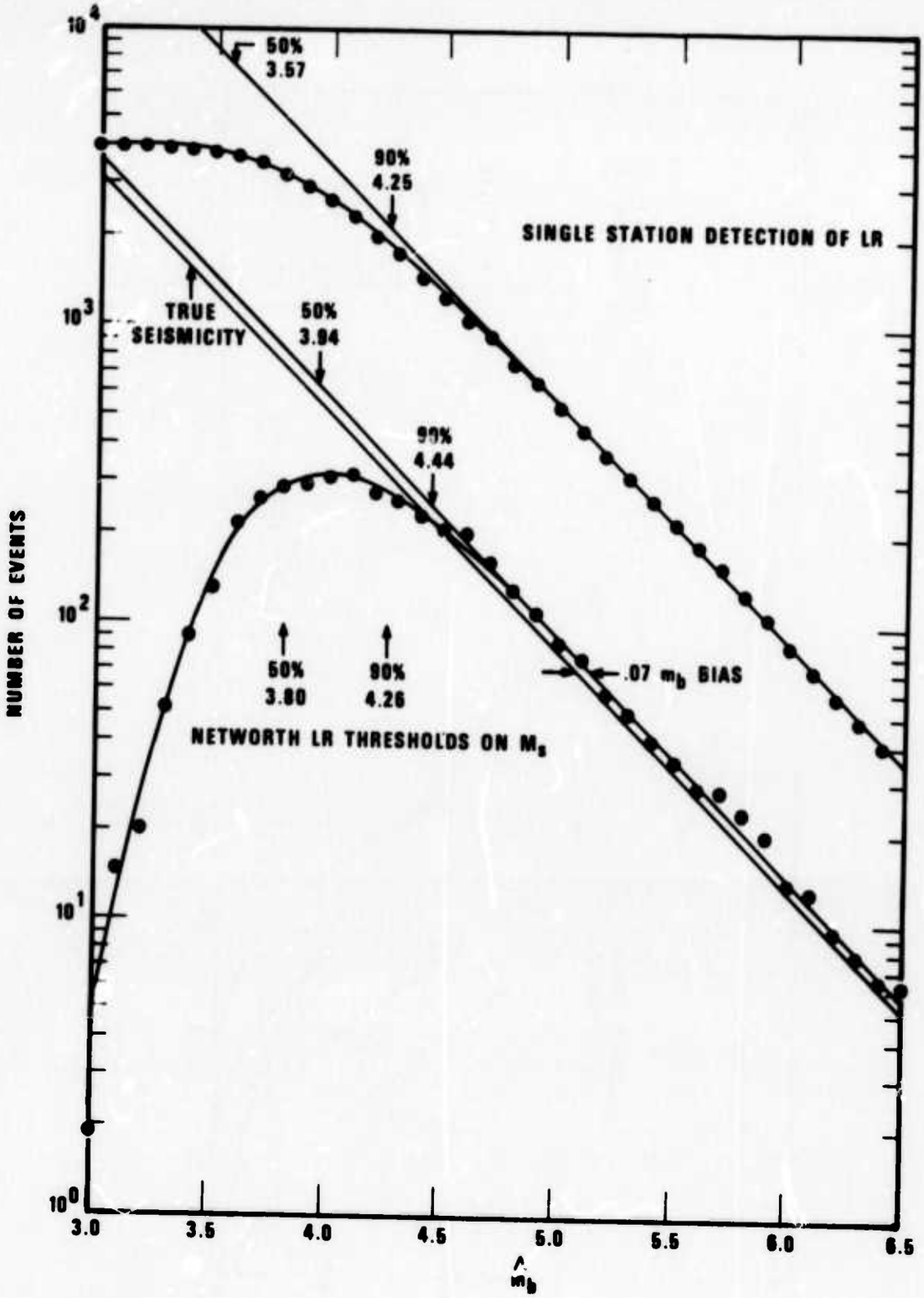


Figure 10. Single-station LR thresholds on $\hat{m}_b$ with a P threshold of 3.47.

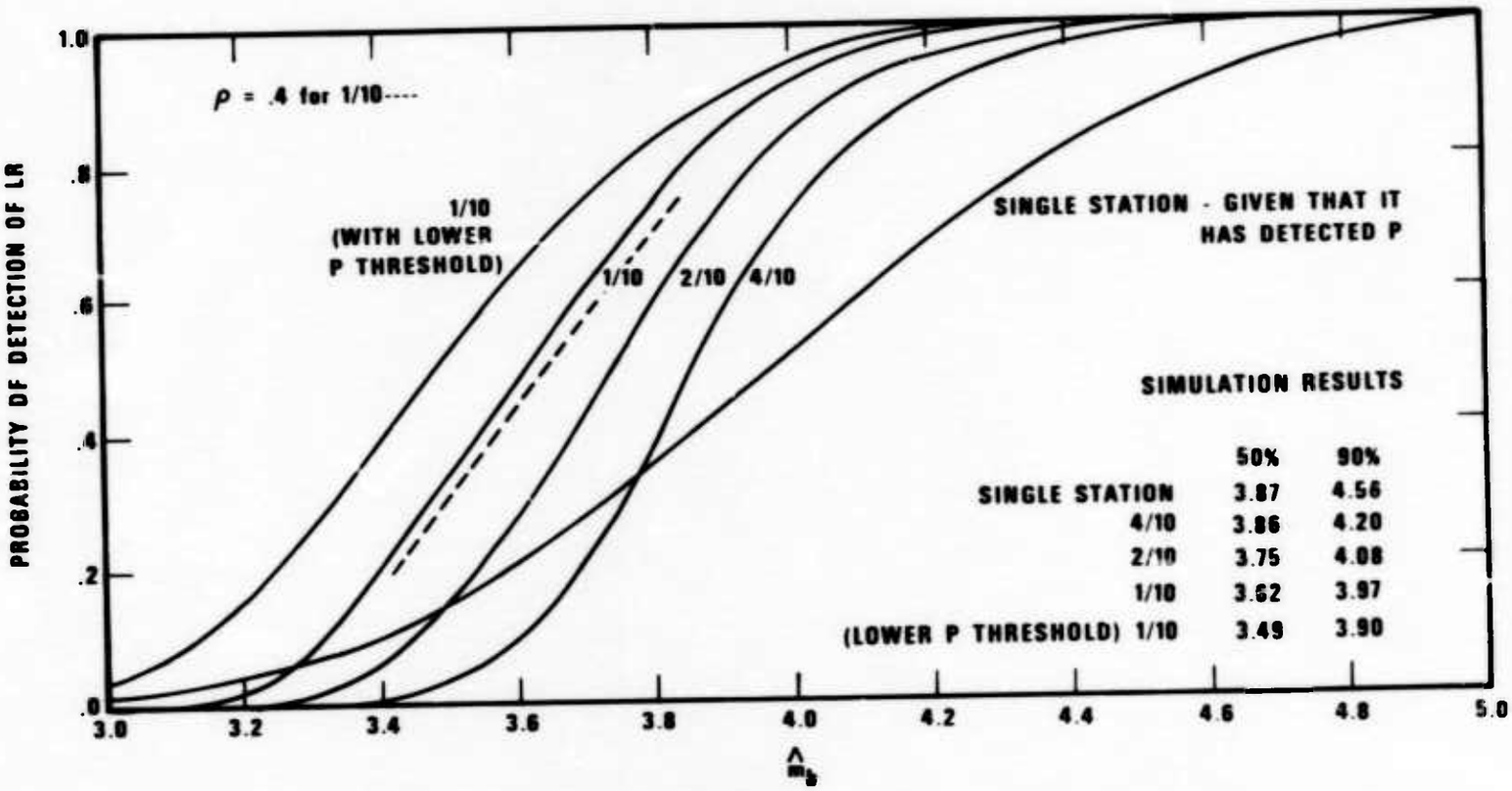


Figure 11. Direct LR thresholds for the simulation network given that >4 station P detection has been declared - source bias present.

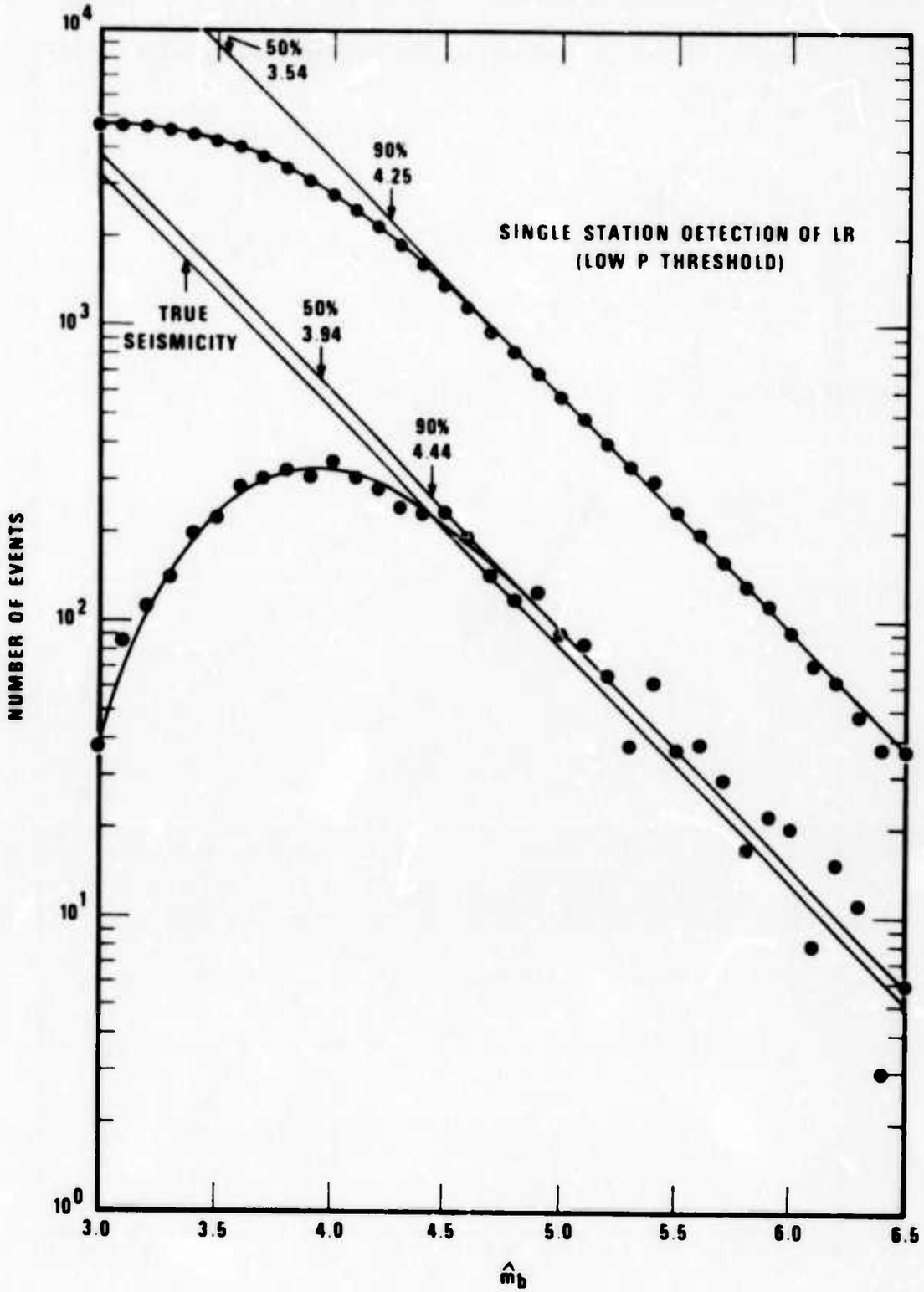


Figure 12. Single-station LR threshold on $\hat{m}_b$ with a P threshold of 3.00 instead of 3.47.

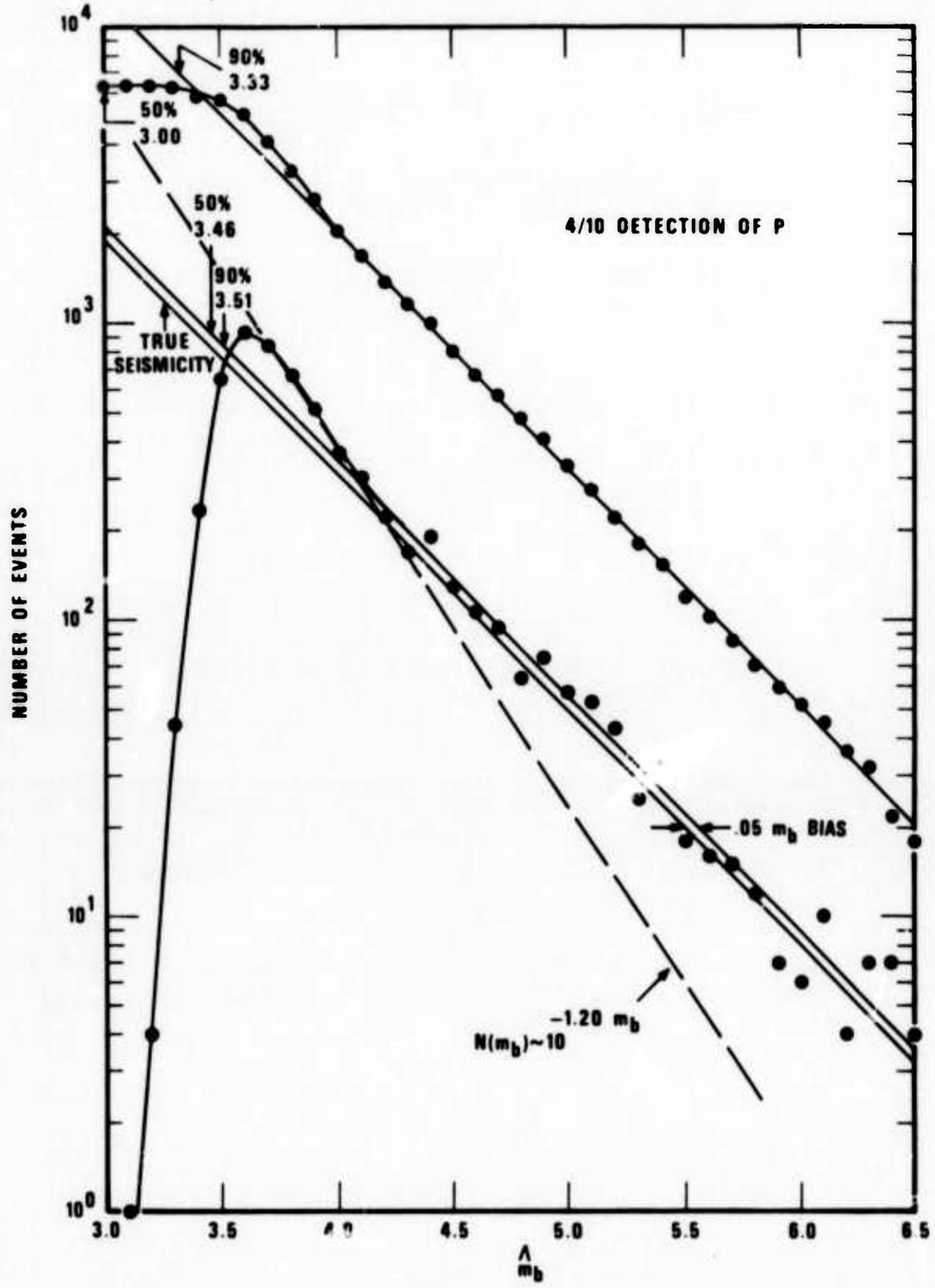


Figure 13. Network P (or LR) thresholds on $\hat{m}_b$ (or $\hat{M}_s$) for $\geq$4 out of 10 detecting with source bias.

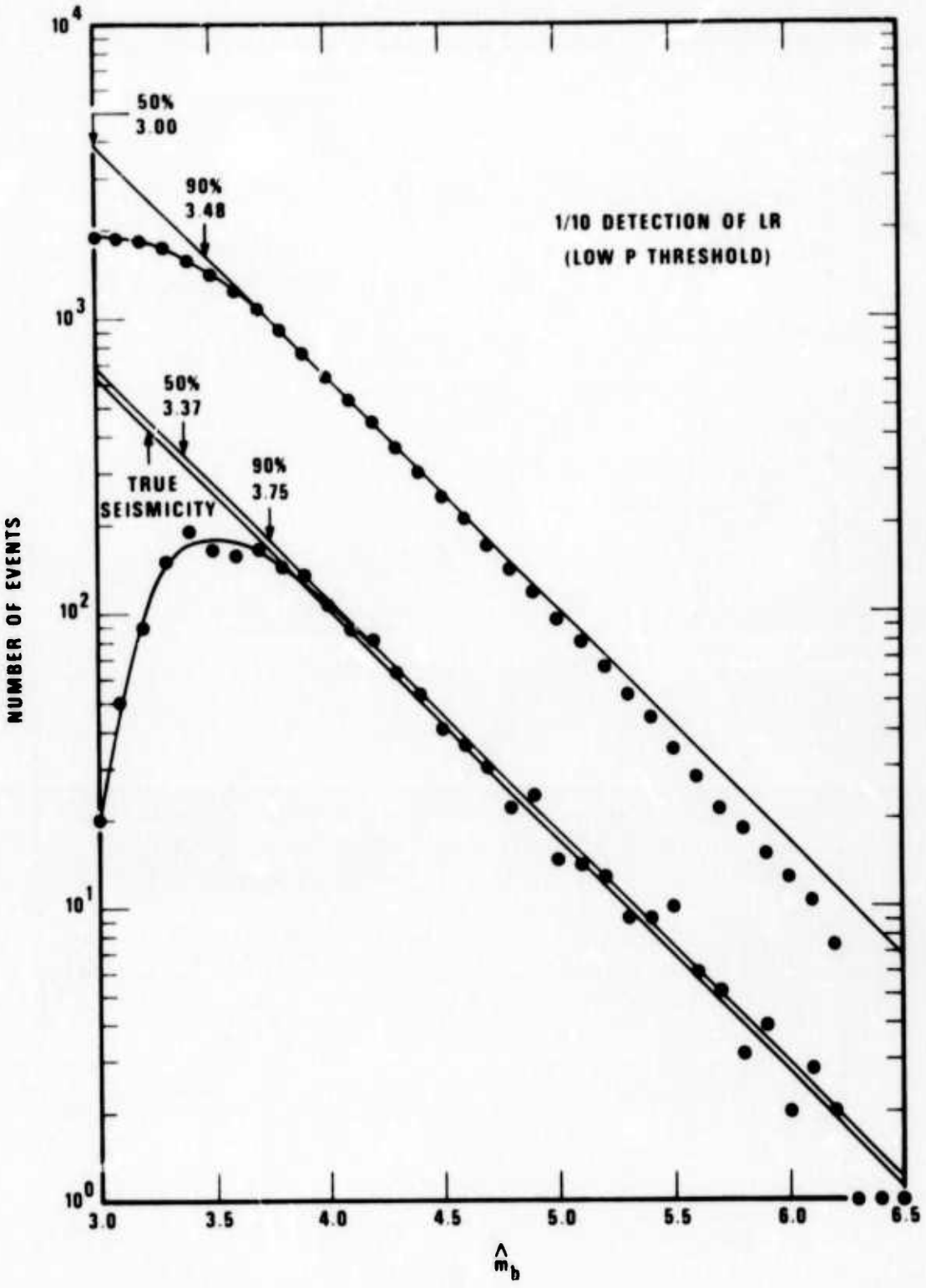


Figure 14. Network LR thresholds on $\hat{m}_b$ for ≥1 out of 10 detecting with source bias and low P threshold (3.00).

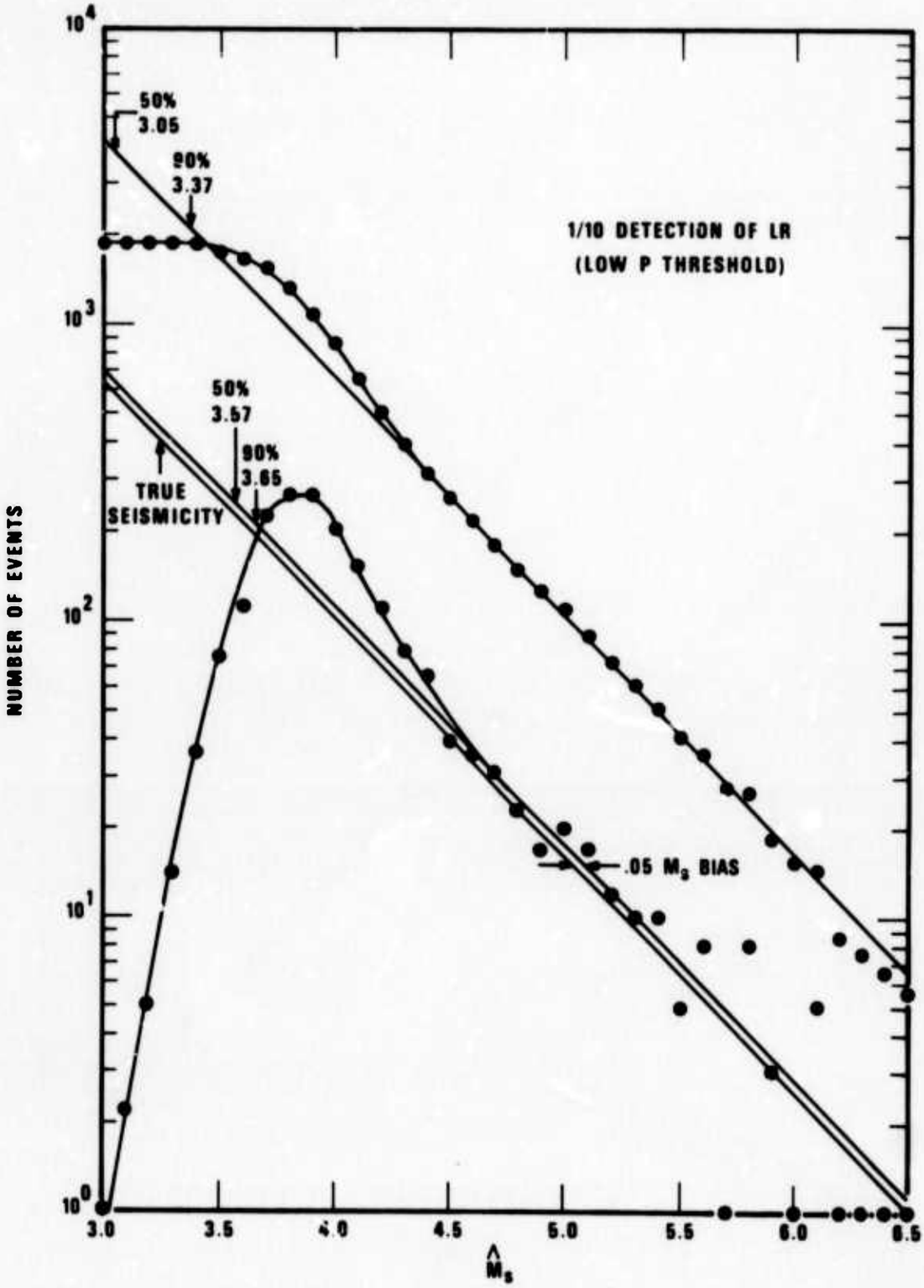


Figure 15. Network LR thresholds on $\hat{M}_s$ for ≥1 out of 10 detecting with source bias and low $\bar{P}$ threshold (3.00).

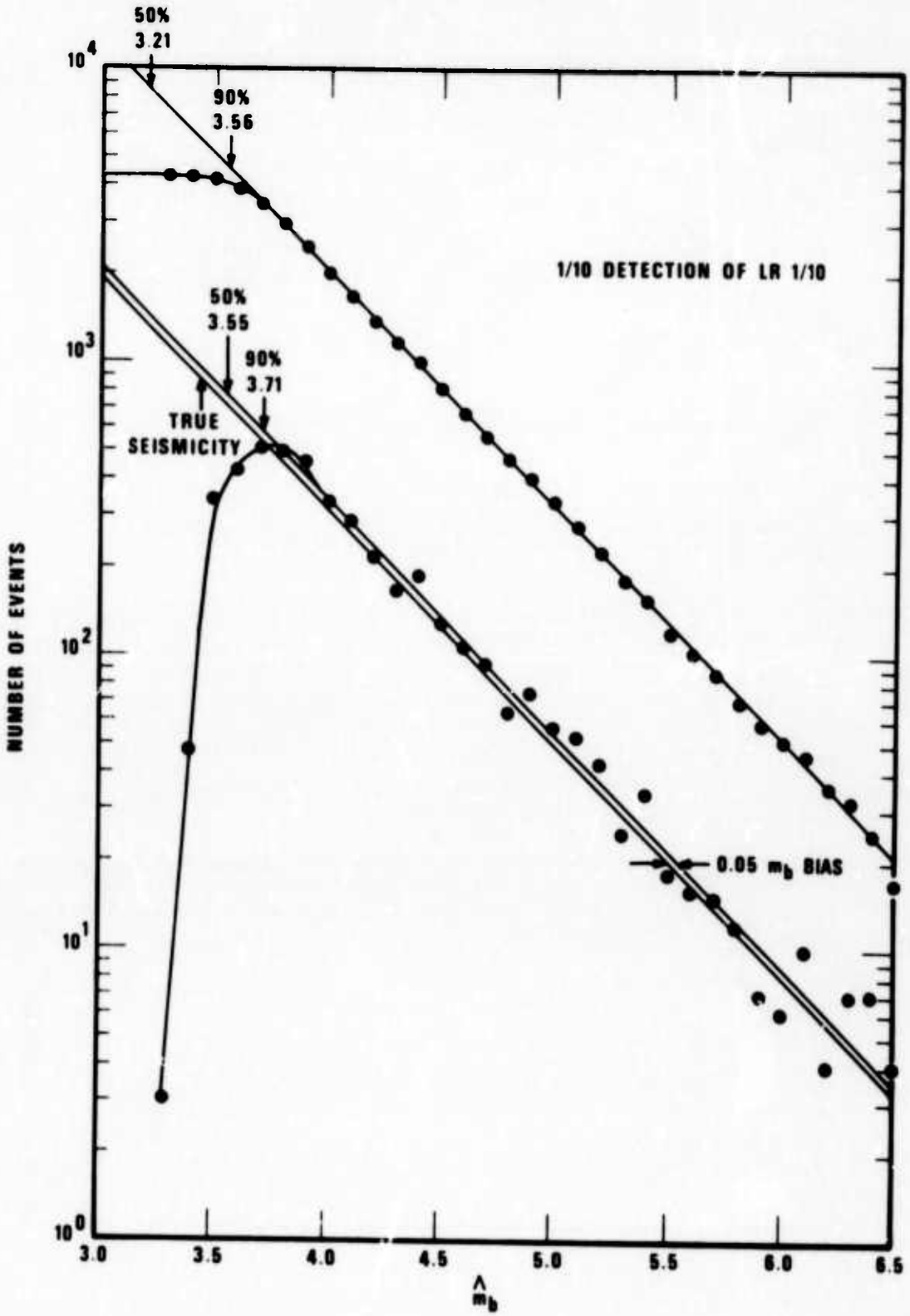


Figure 16a. Network LR thresholds on $\hat{m}_b$ for $\geq$1 out of 10 detecting with source bias and standard P threshold (3.47).

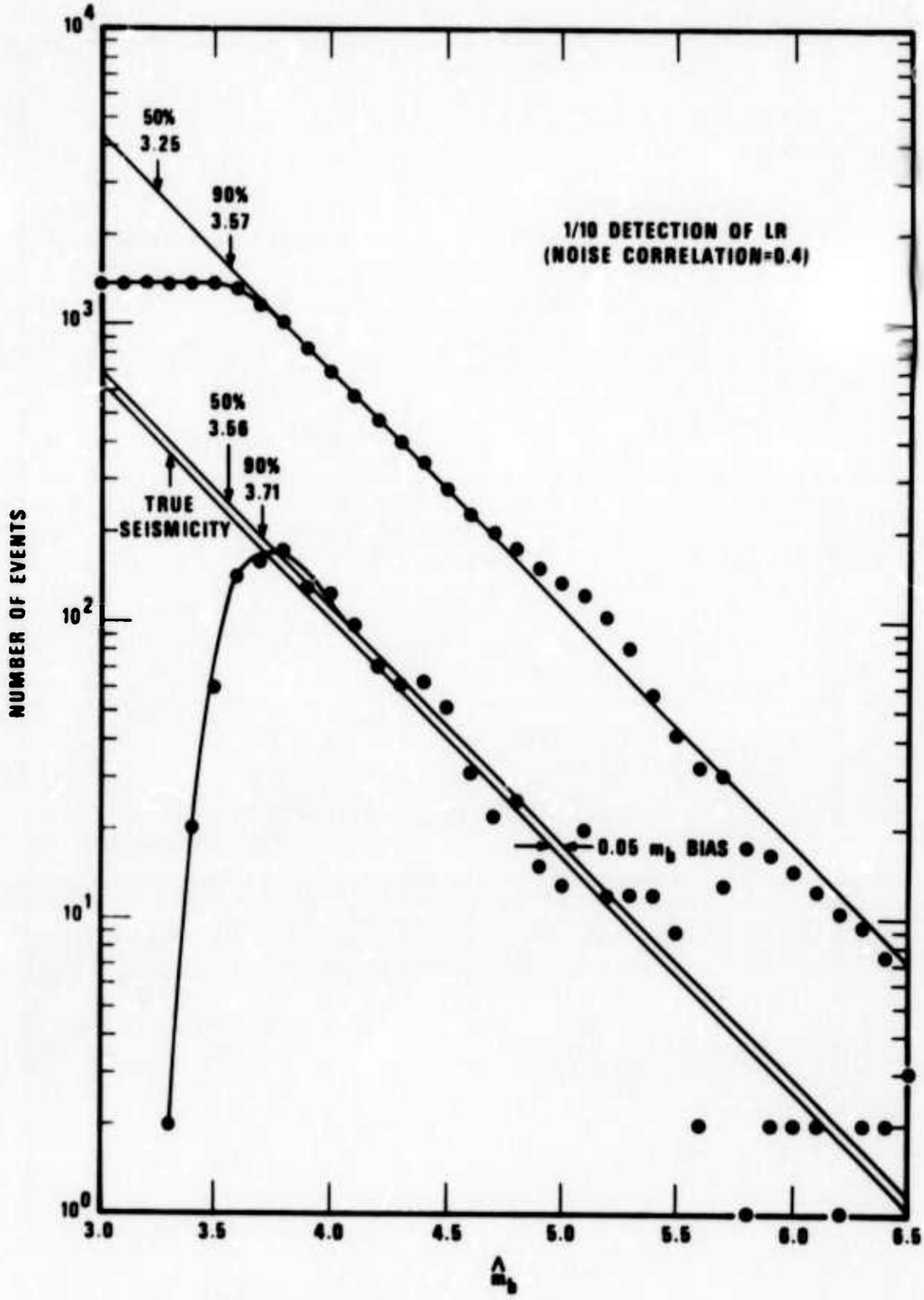


Figure 16b. Network LR thresholds on m_b for ≥1 out of 10 detection with source bias and standard P threshold (3.47) and correlated noise (ρ_n=.4).

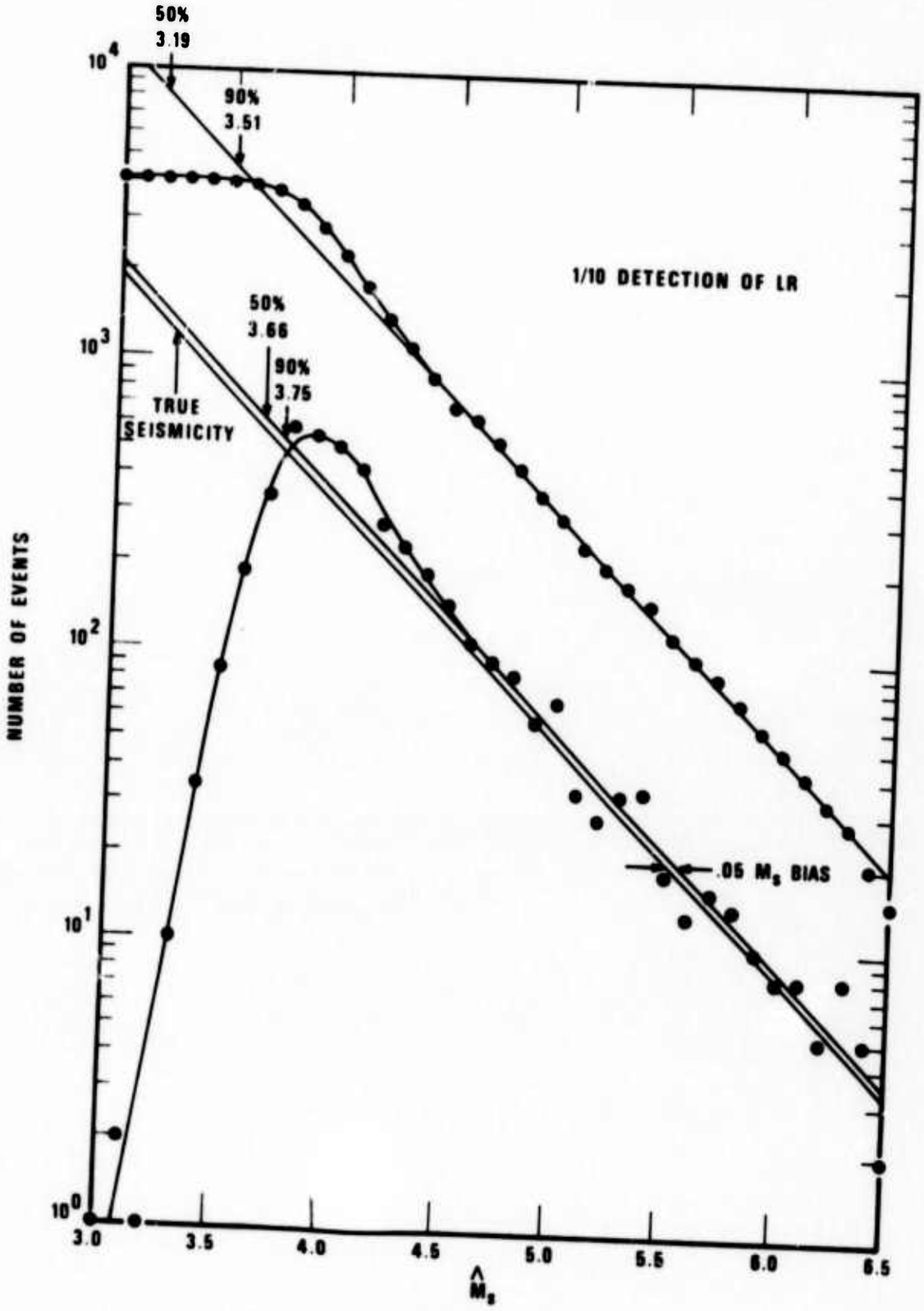


Figure 17a. Network LR thresholds on $\hat{M}_s$ for $\geq$1 out of 10 detecting with source bias and standard P threshold (3.47).

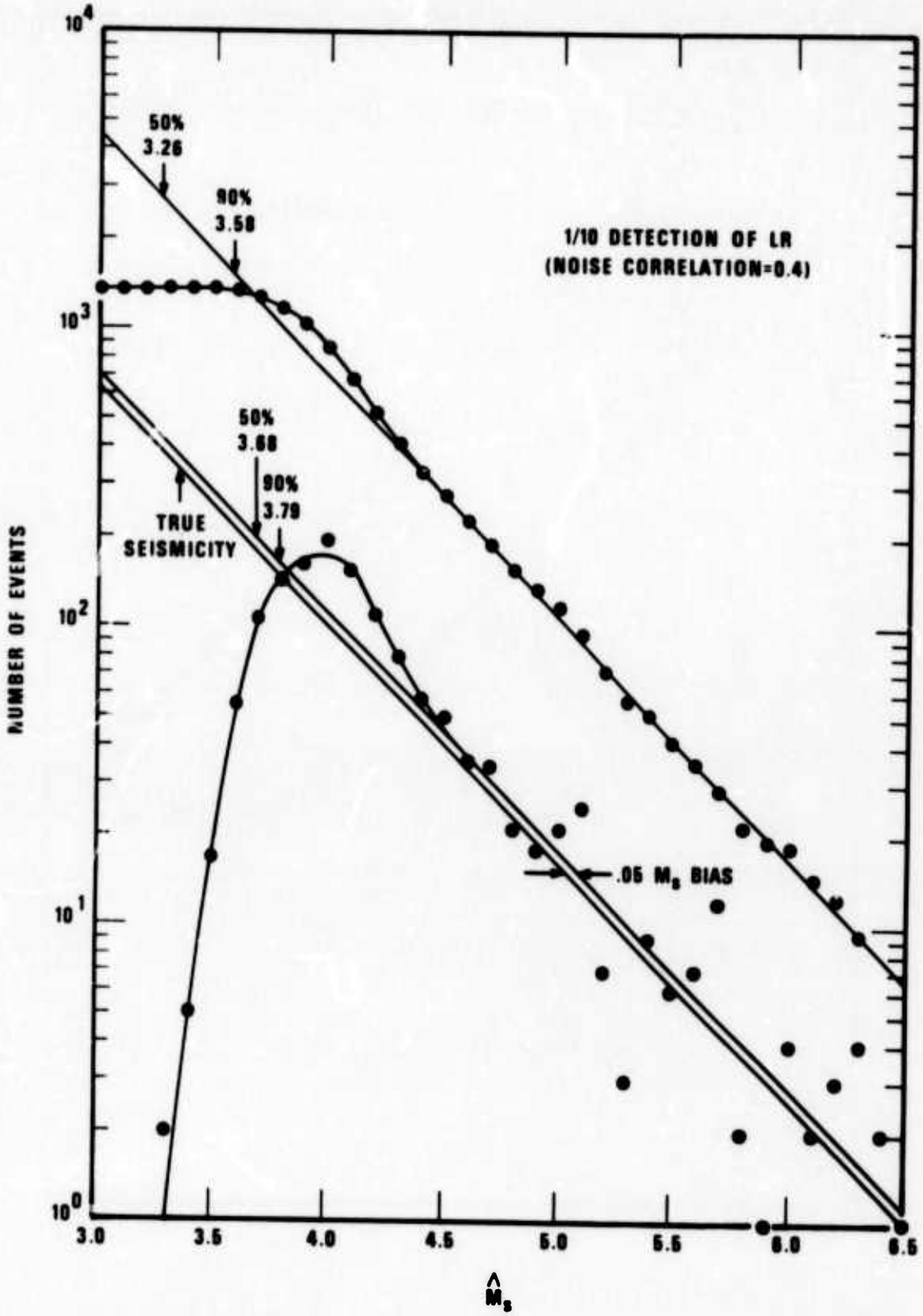


Figure 17b. Network LR thresholds on $\hat{M}_s$ for $\geq$1 out of 10 detecting with source bias and standard P threshold (3.47) and correlated noise (ρ_n=.4).

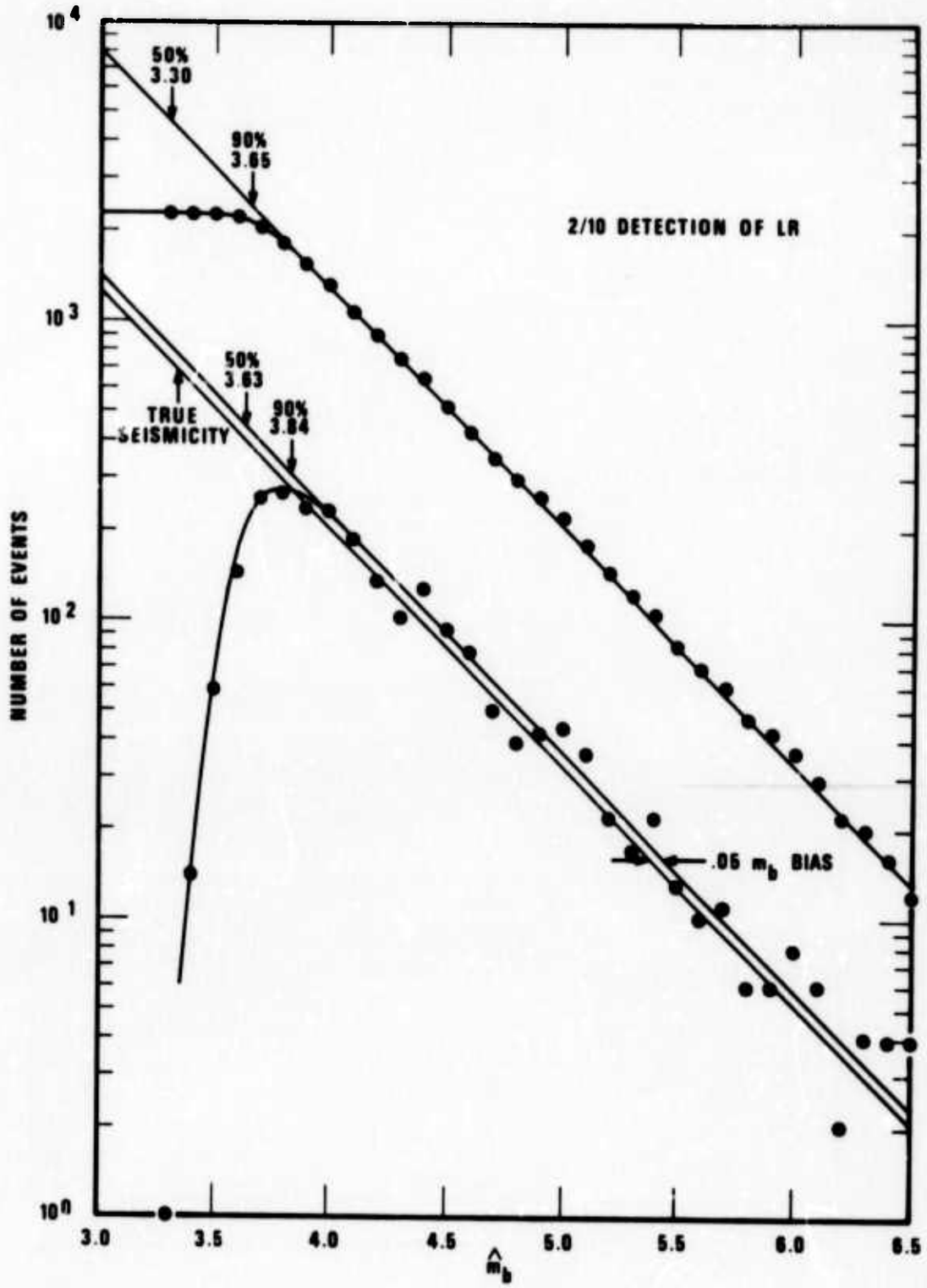


Figure 18. Network LR thresholds on $\hat{m}_b$ for $\geq$2 out of 10 detecting with source bias and standard P threshold (3.47).

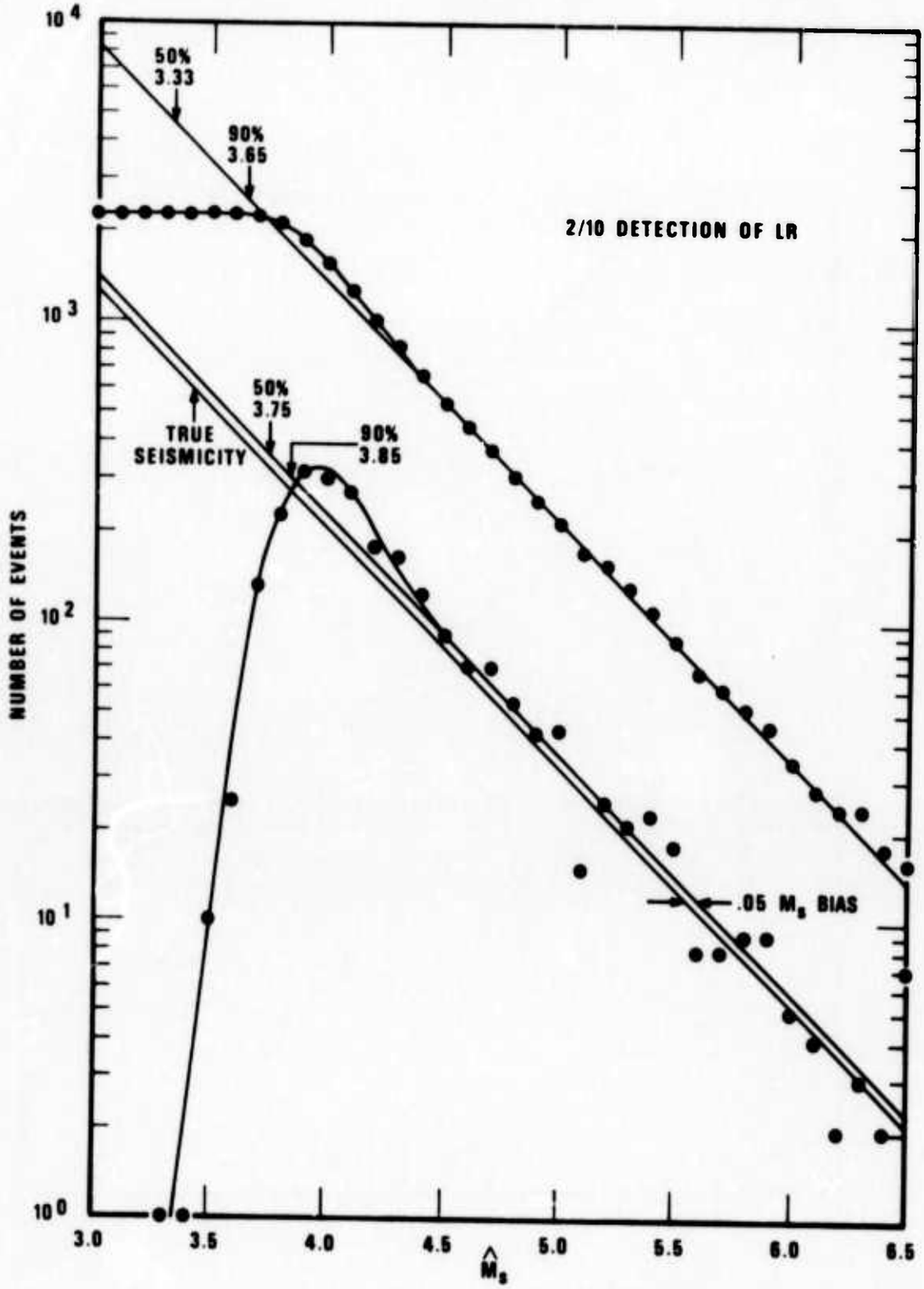


Figure 19. Network LR thresholds on $\hat{M}_s$ for $\geq$2 out of 10 detecting with source bias and standard P threshold (3.47).

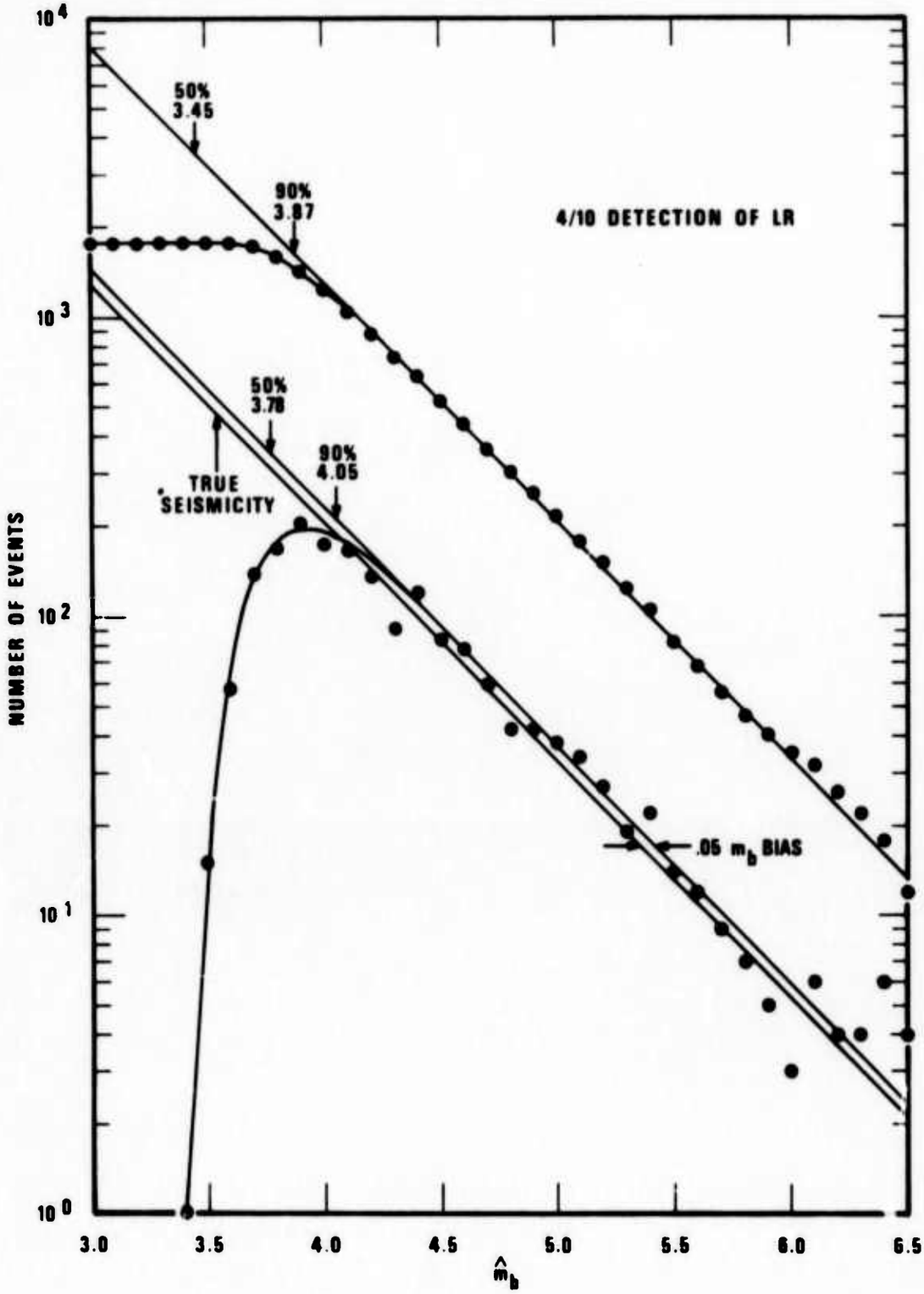


Figure 20. Network LR thresholds on $\hat{m}_b$ for $\geq$4 out of 10 detecting with source bias and standard P threshold (3.47).

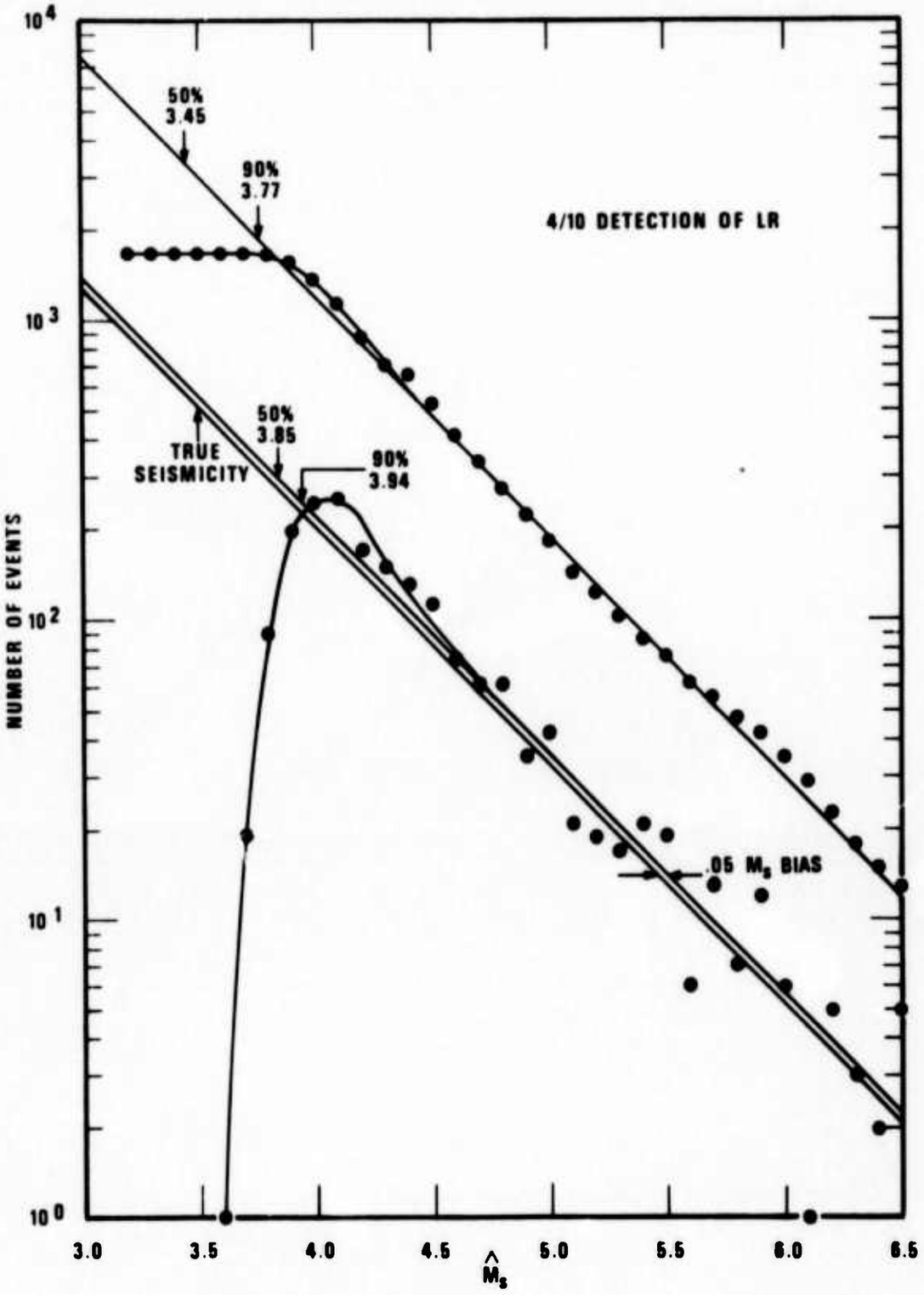


Figure 21. Network LR thresholds on $\hat{M}_s$ for ≥4 out of 10 detecting with source bias and standard P threshold (3.47).

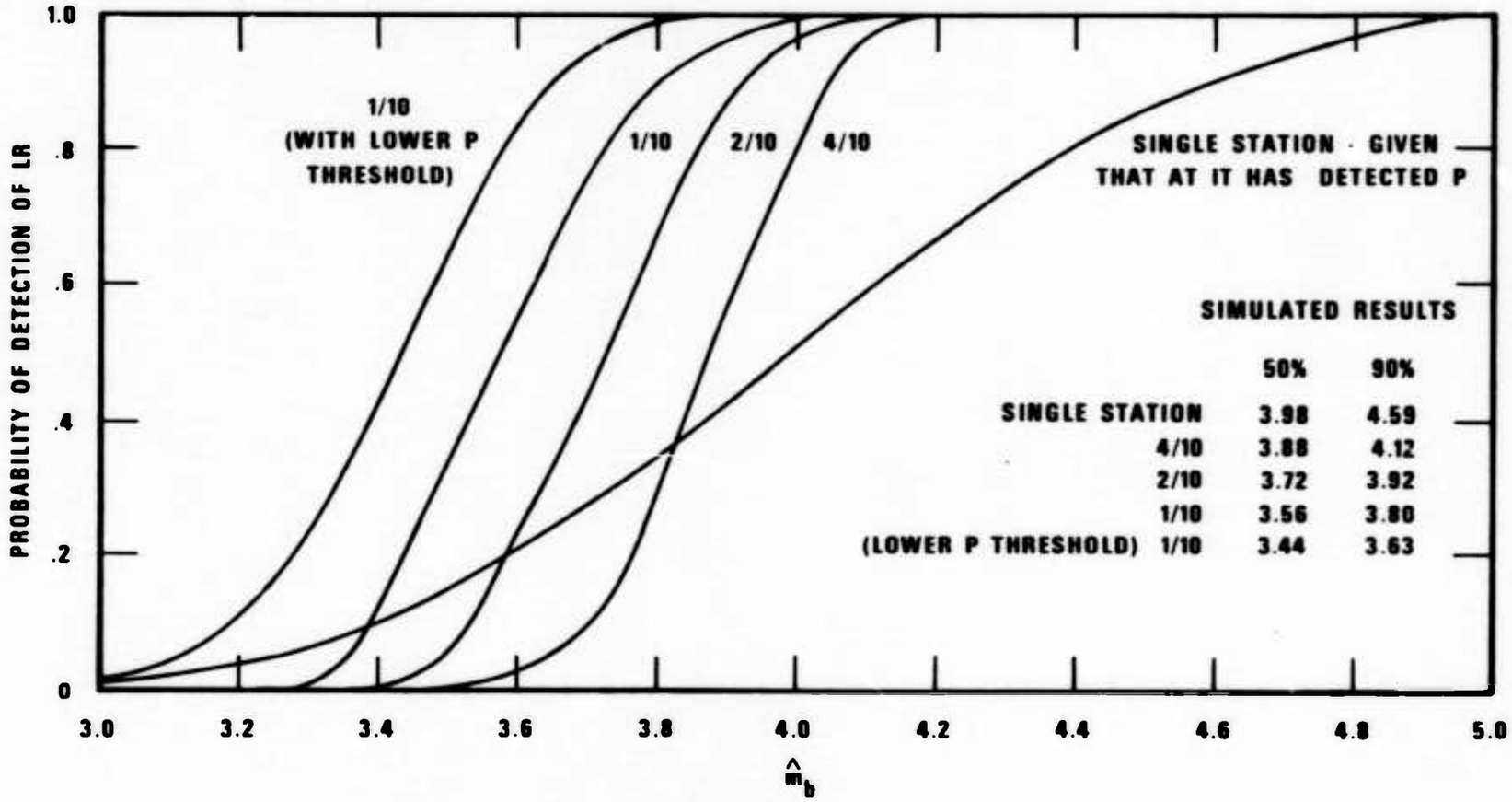


Figure 22. Direct LR thresholds for the simulation network given that a ≥ 4 station P detection has been declared - no source bias present.

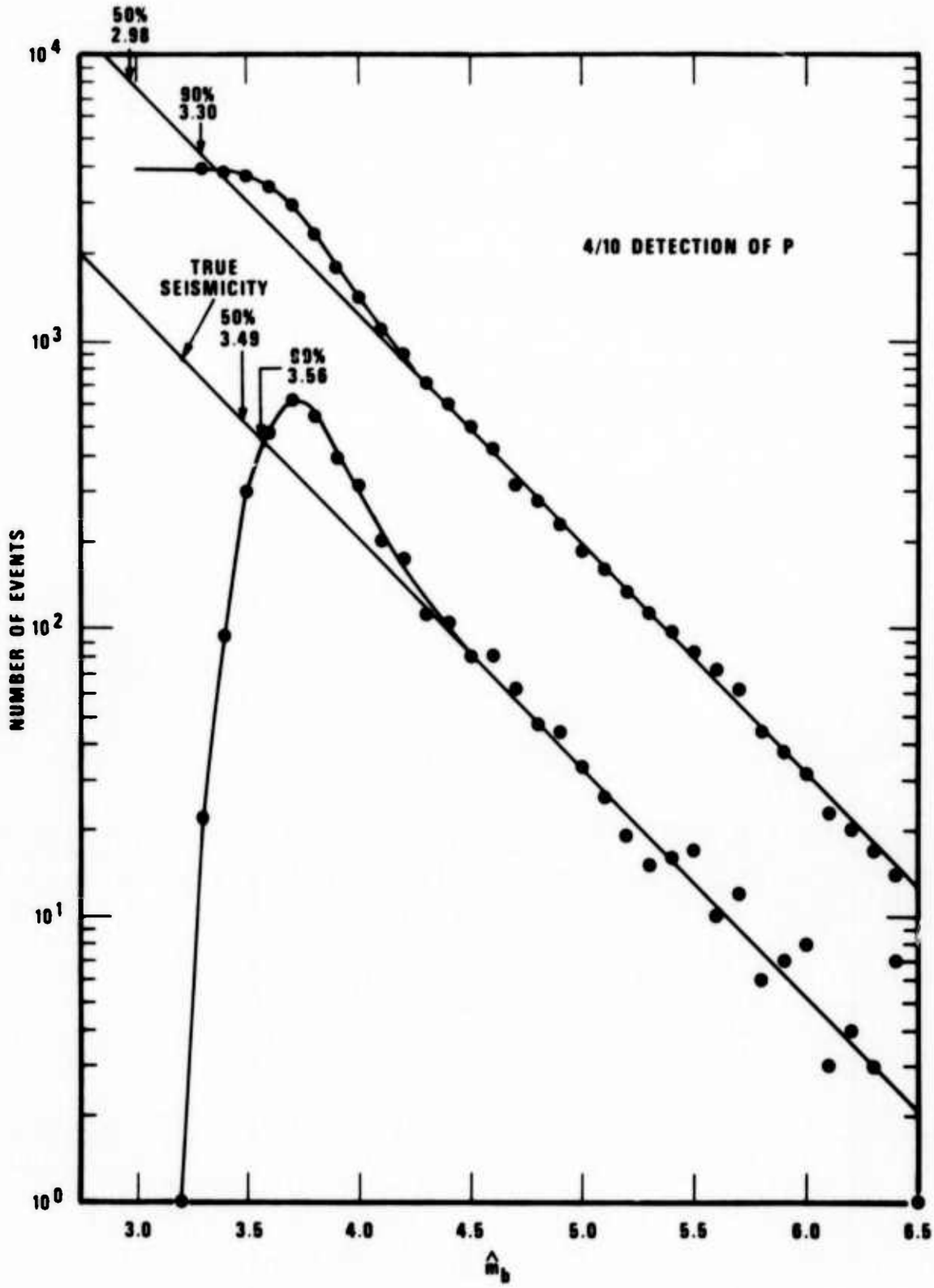


Figure 23. Network P (or LR) thresholds on $\hat{m}_b$ (or $\hat{M}_s$) for $\geq$4 out of 10 detecting without source bias.

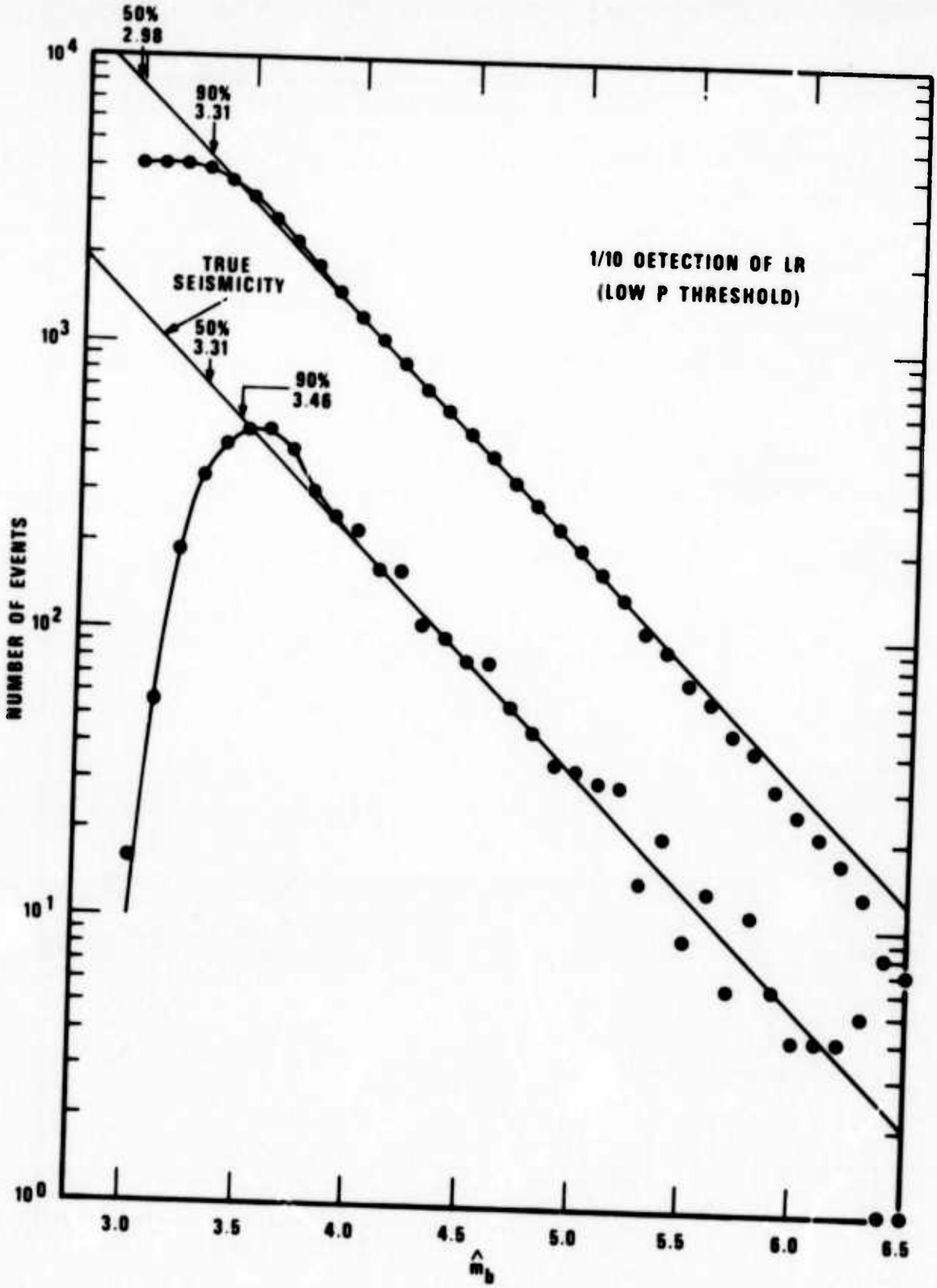


Figure 24. Network LR thresholds on $\hat{m}_b$ for ≥ 1 out of 10 detecting without source bias and with low $\bar{P}$ threshold (3.00).

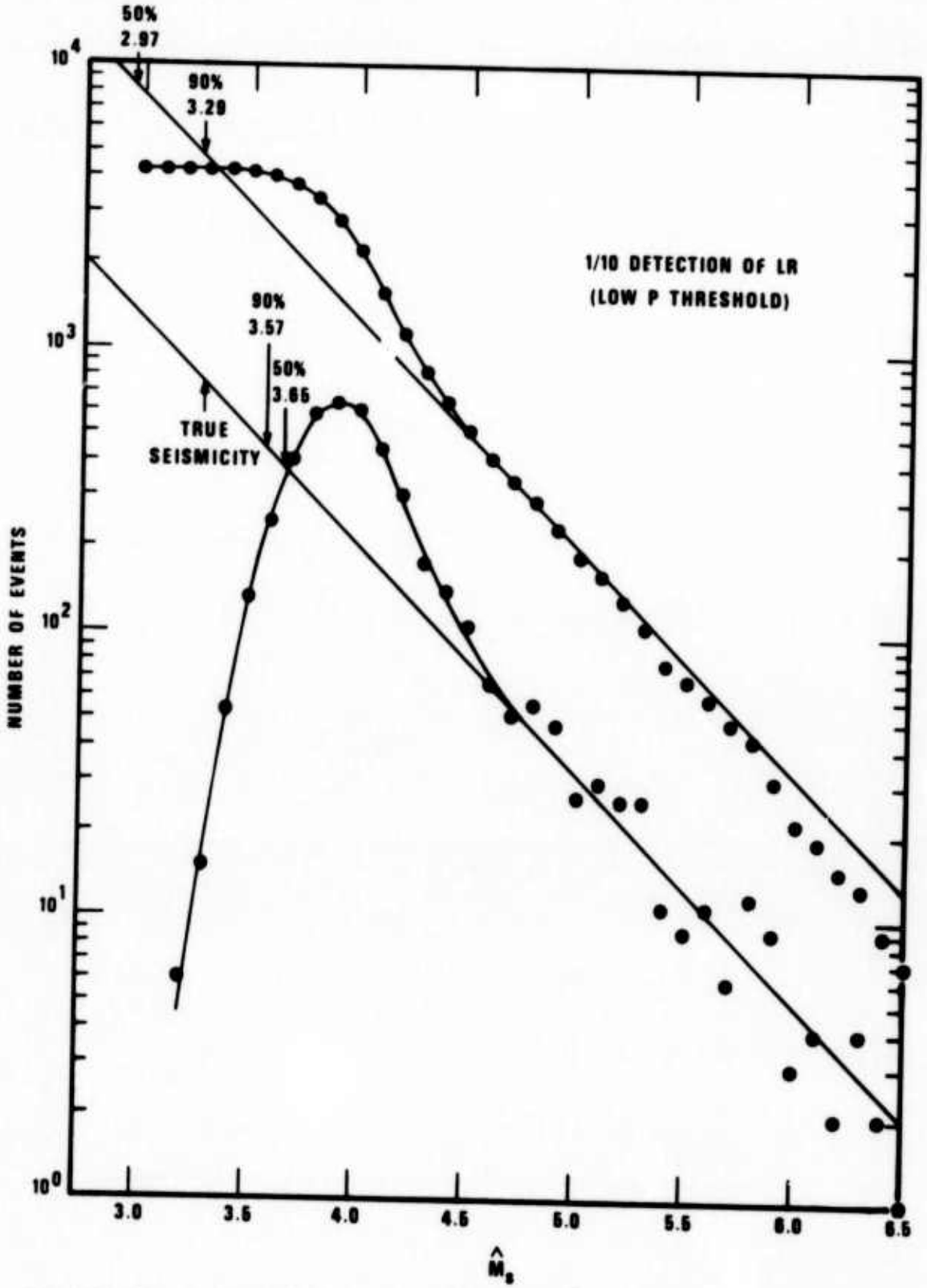


Figure 25. Network LR thresholds on $\hat{M}_s$ for $\geq$1 out of 10 detecting without source bias and with low $\bar{P}$ threshold (3.00).

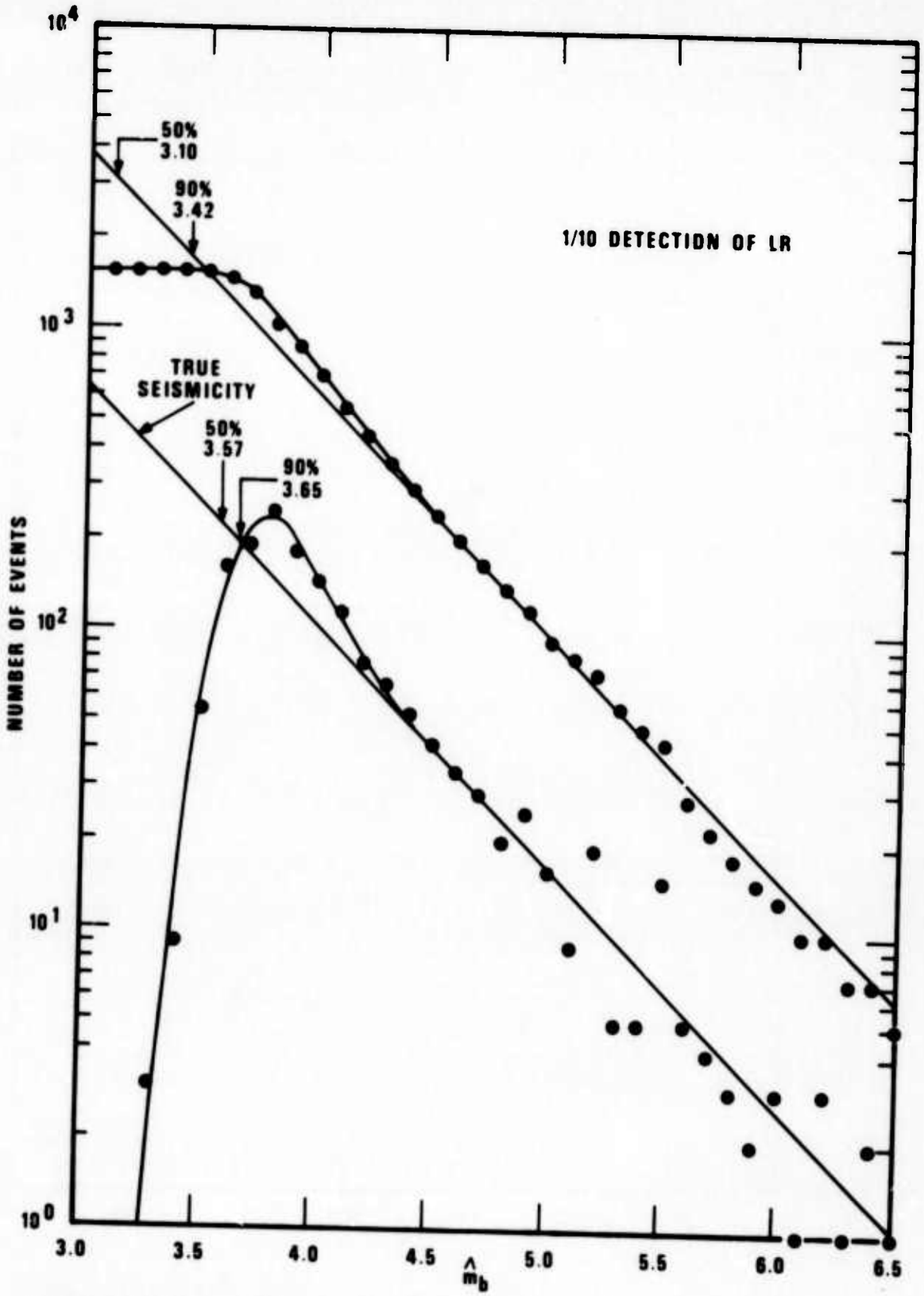


Figure 26. Network LR thresholds on $\hat{m}_b$ for $\geq$1 out of 10 detecting without source bias and with standard P threshold (3.47).

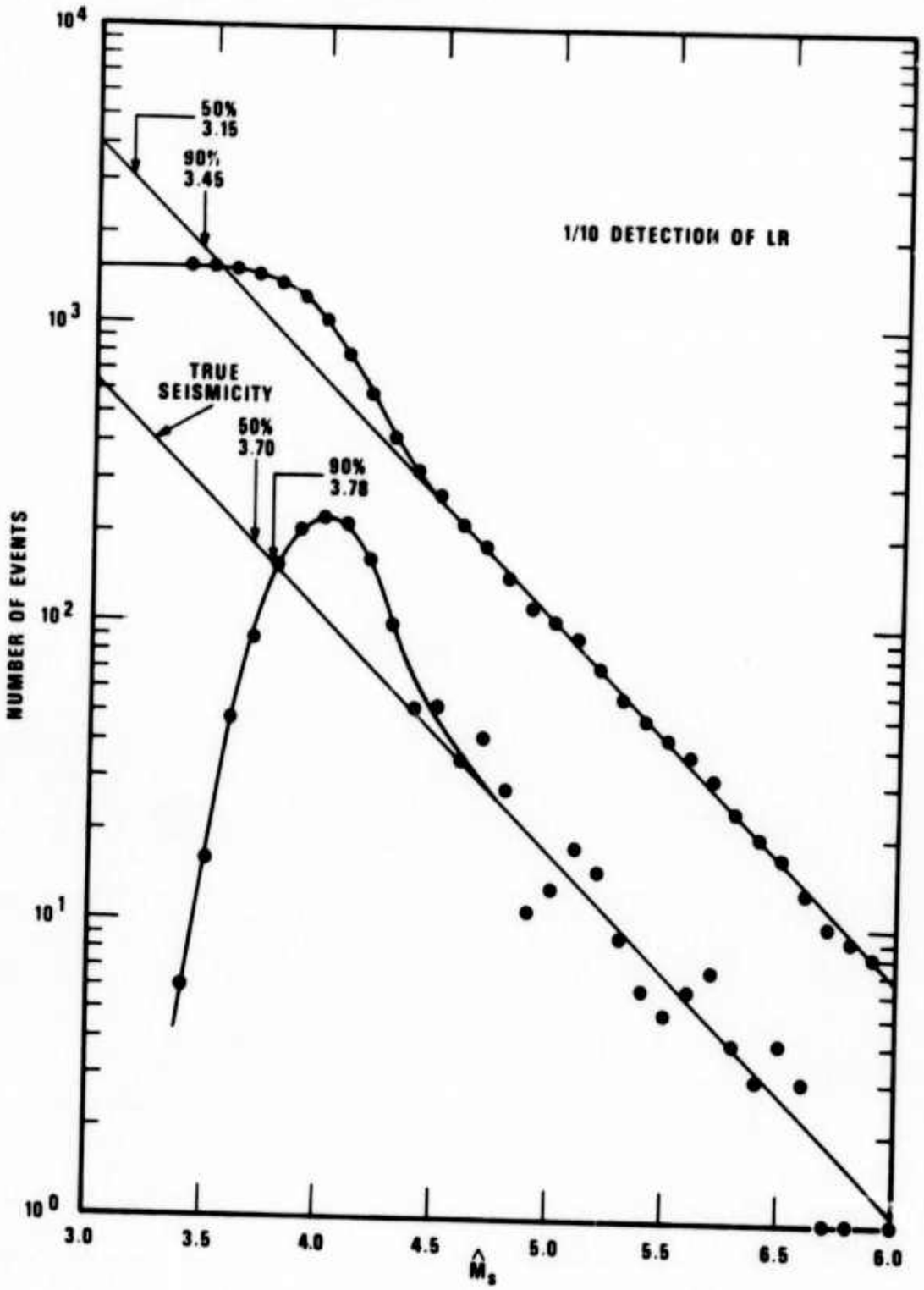


Figure 27. Network LR thresholds on $\hat{M}_s$ for $\geq$1 out of 10 detecting without source bias and with standard P threshold (3.47).

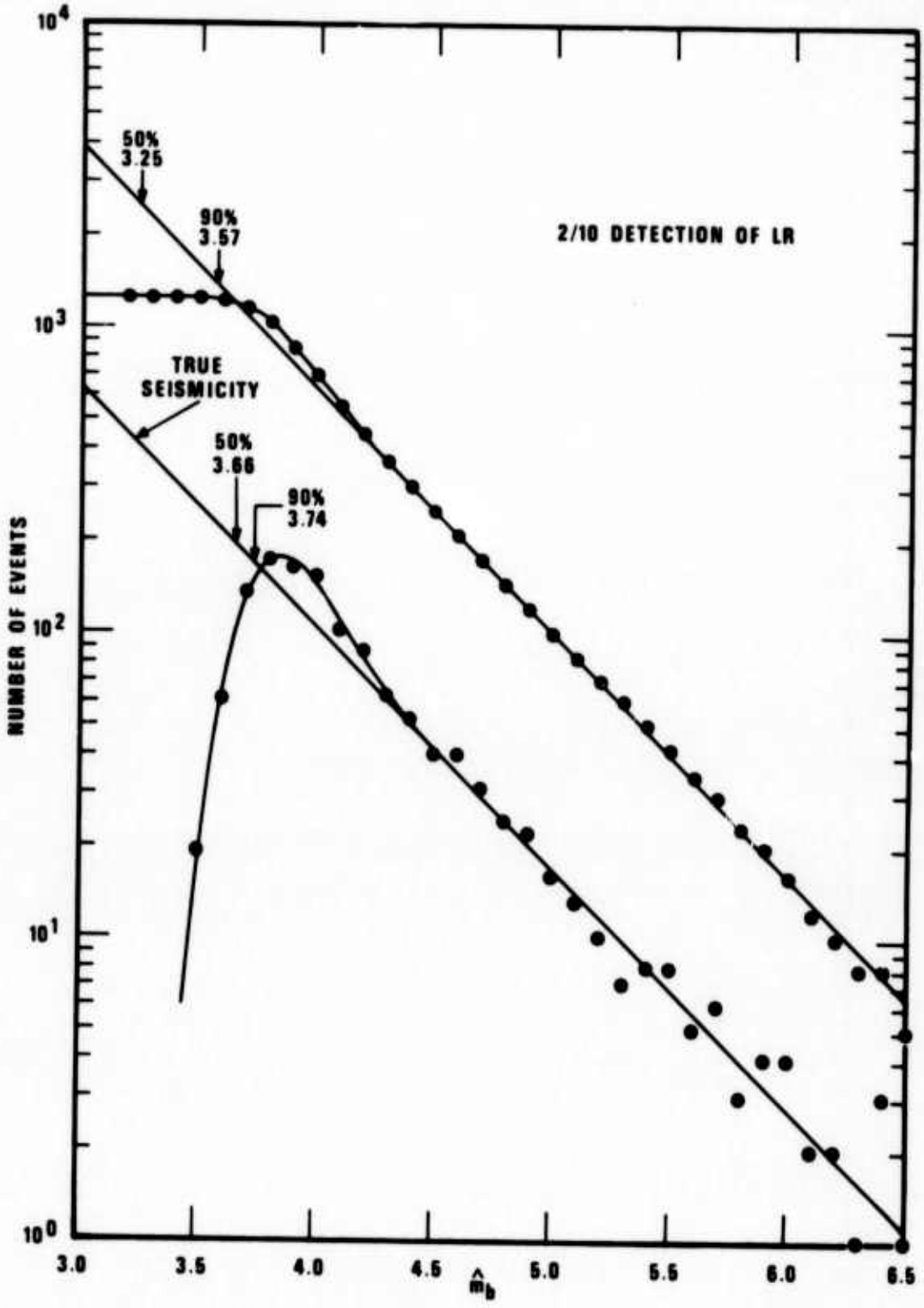


Figure 28. Network LR thresholds on $\hat{m}_b$ for ≥2 out of 10 detecting without source bias and with standard P threshold (3.47).

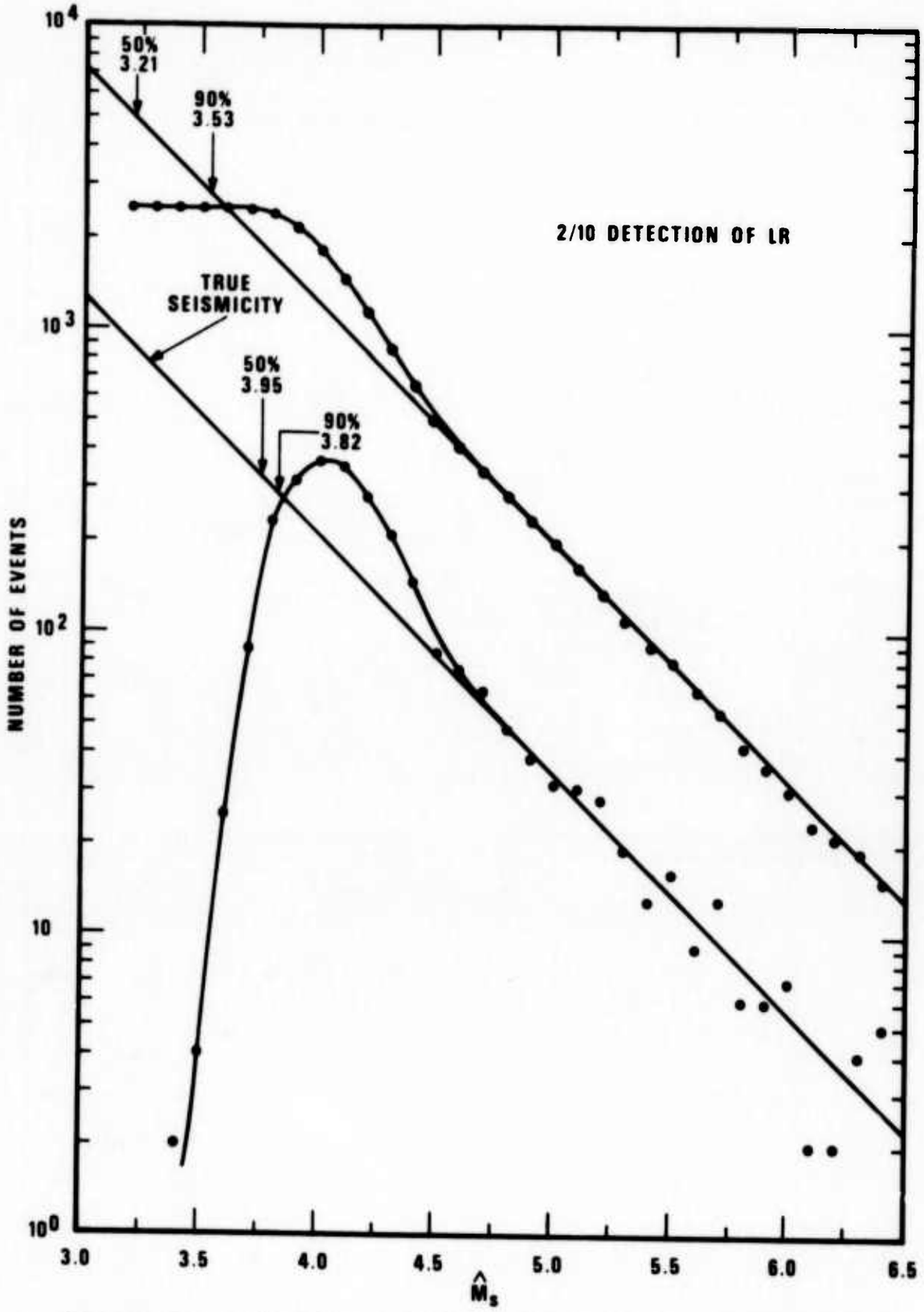


Figure 29. Network LR thresholds on M_S for ≥2 out of 10 detecting without source bias and with standard P threshold (3.47).

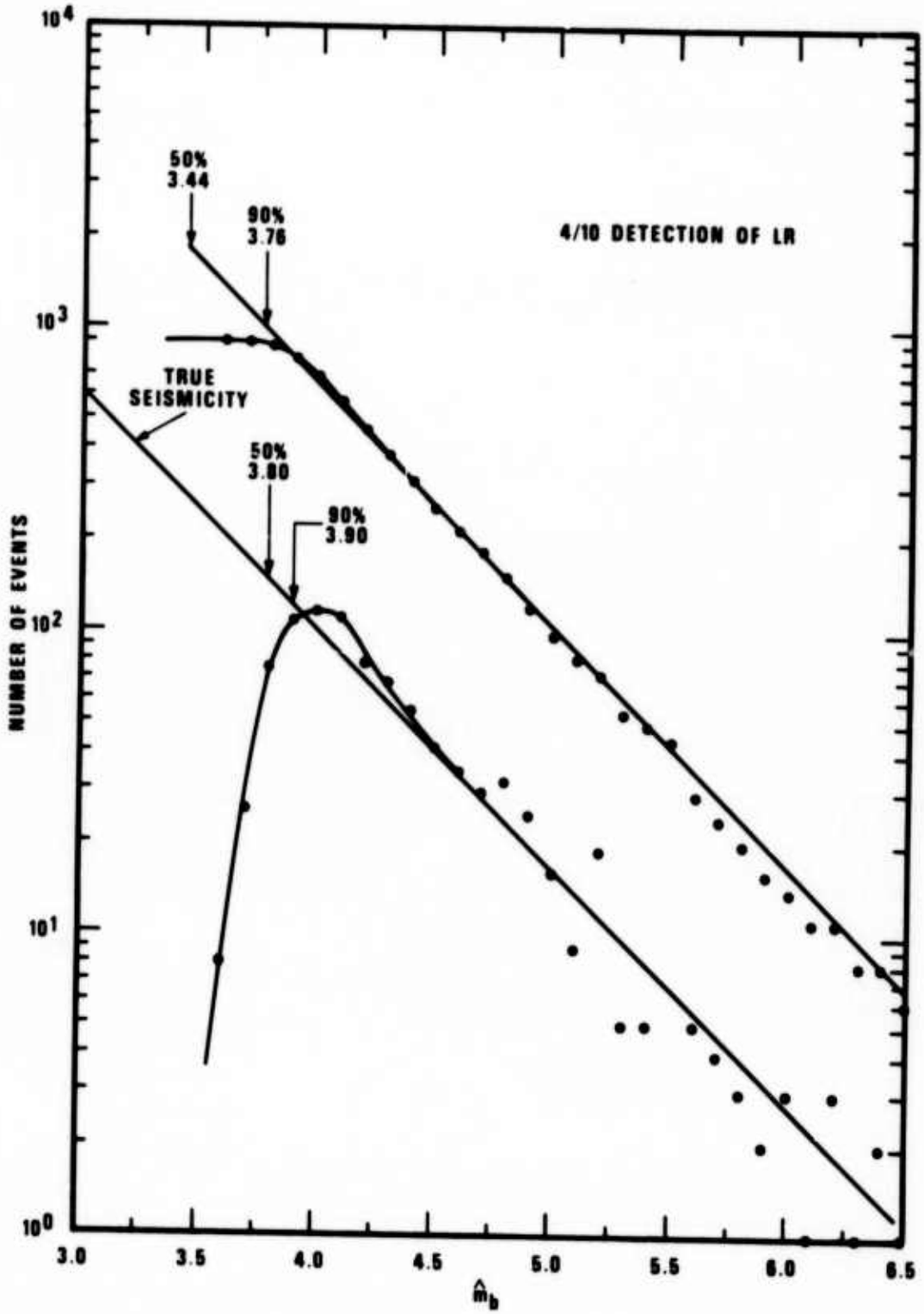


Figure 30. Network LR thresholds on $\hat{m}_b$ for ≥ 4 out of 10 detecting without source bias and with standard P threshold (3.47).

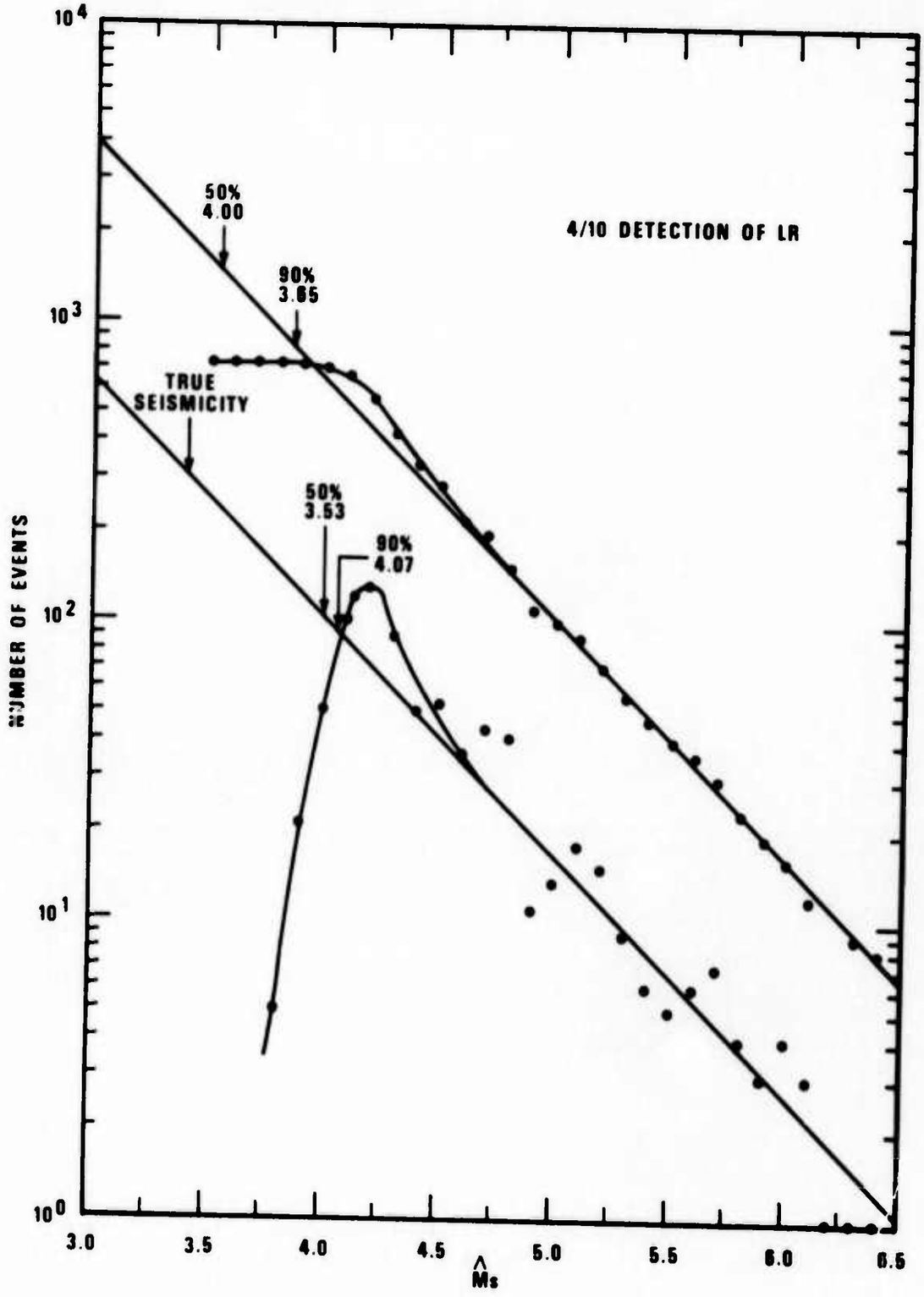


Figure 31. Network LR thresholds on $\hat{M}_s$ for ≥4 out of 10 detecting without source bias and with standard P threshold (3.47).

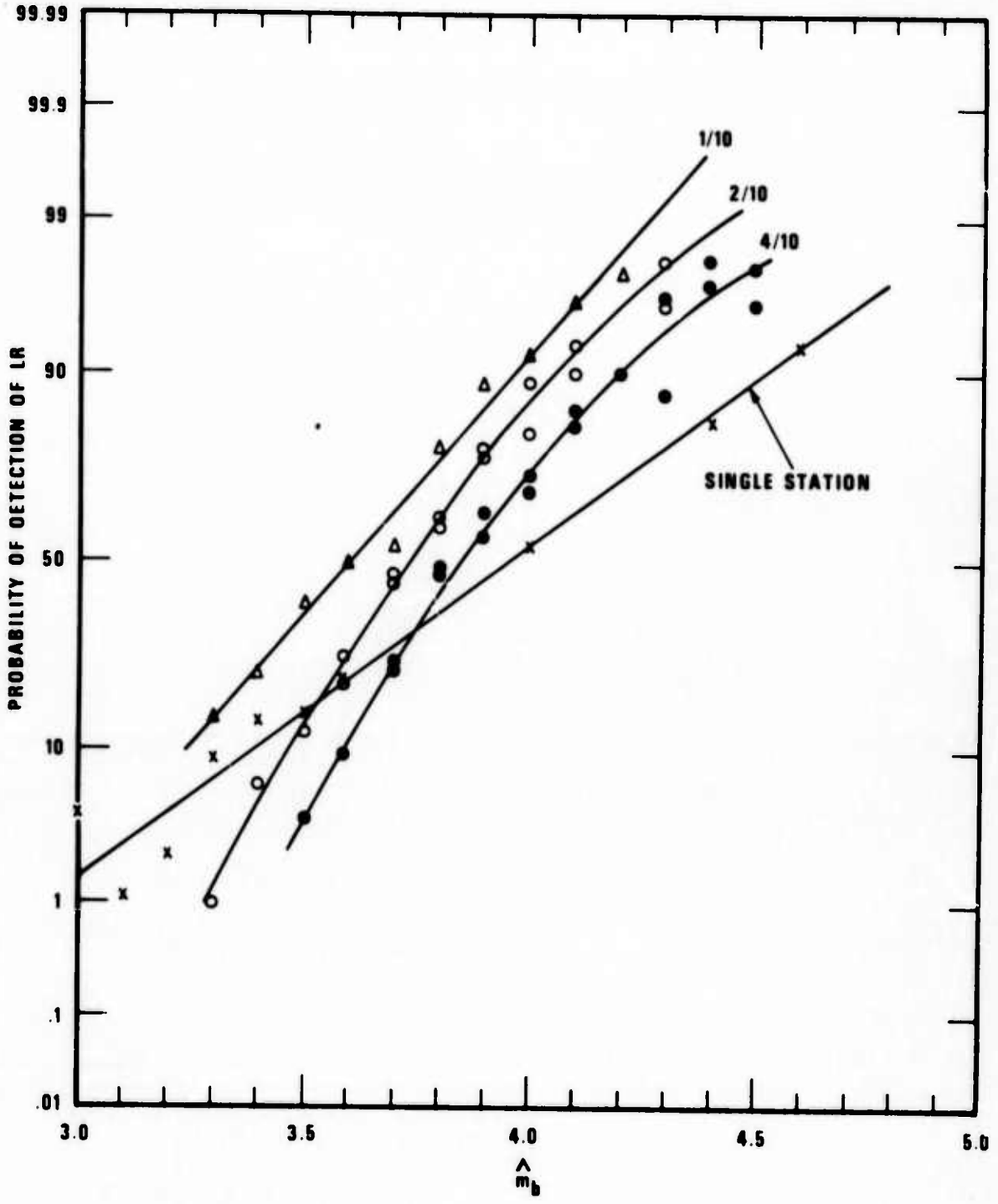


Figure 32. Network LR probability of detection curves on $\hat{m}_b$.

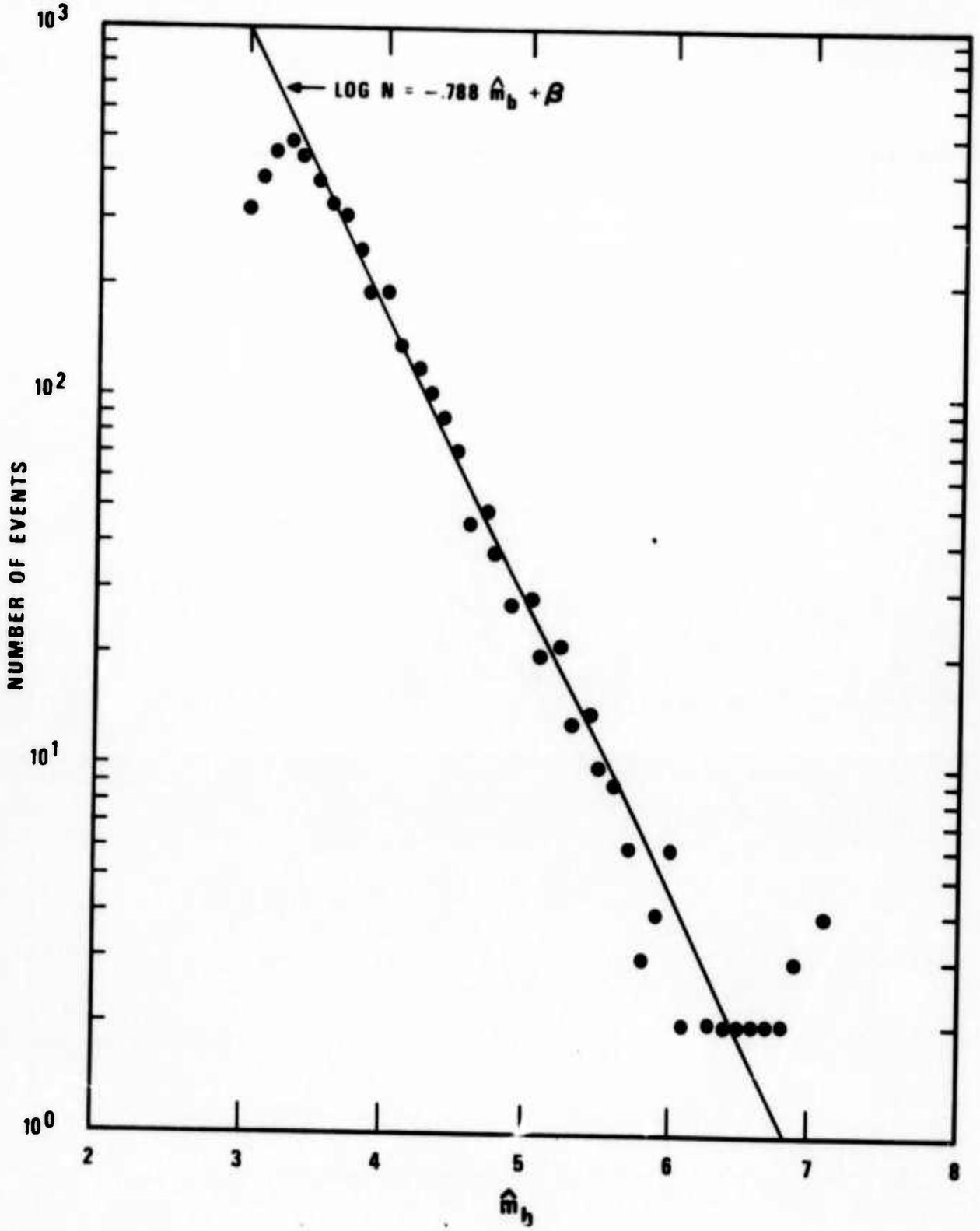


Figure 33. Single-station P thresholds on $\hat{m}_b$ with source bias and with low P threshold (3.00).

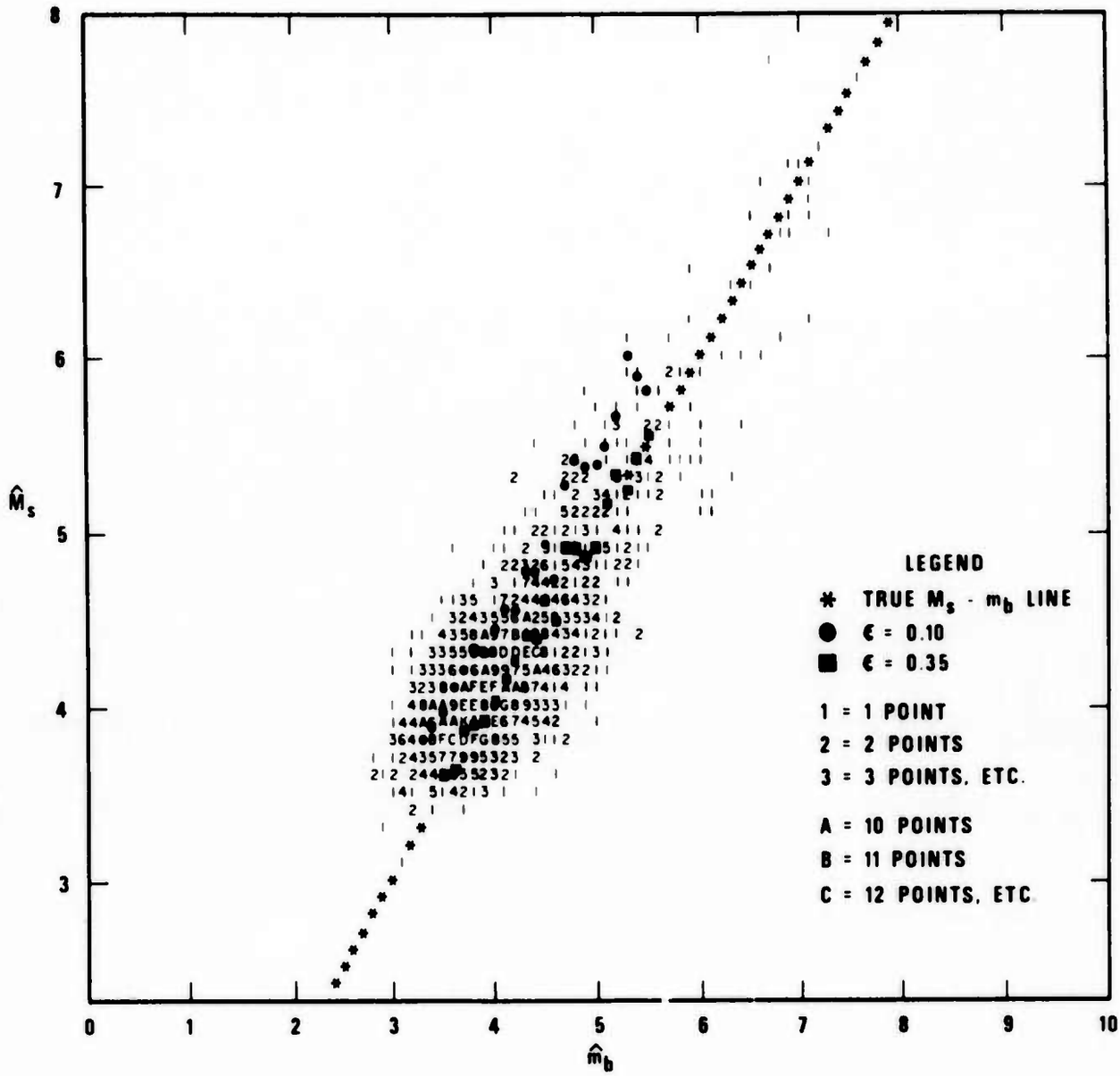


Figure 34. MSBNET simulated (M_s,m_b) points for a single station with 50% P threshold of 3.00 and 50% LR threshold of 3.80.

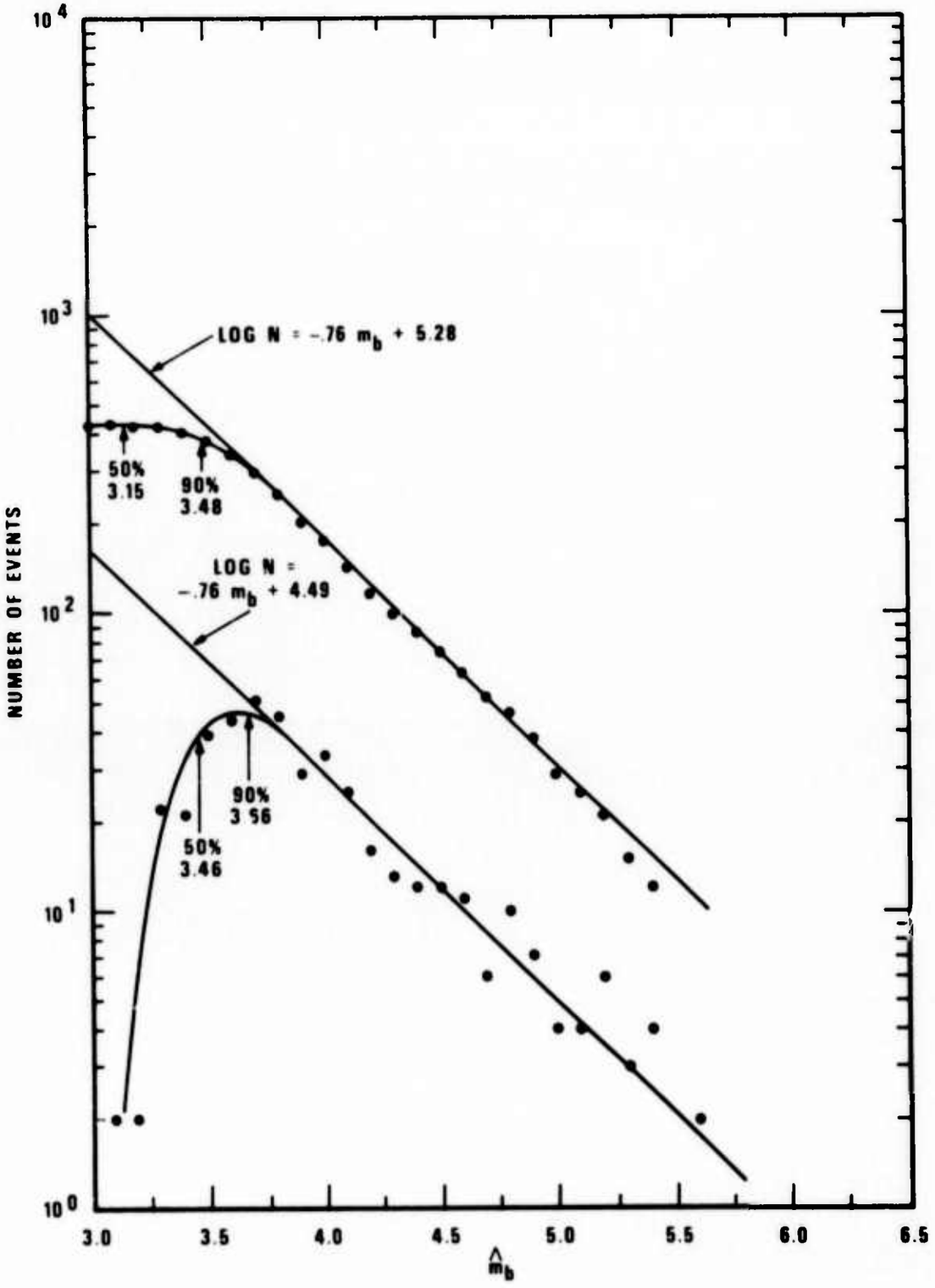


Figure 35. LASA incremental and cumulative thresholds for P waves from events in Kuril-Kamchatka for the time intervals 16 February to 5 March 1972 and 30 April to 30 September 1972.

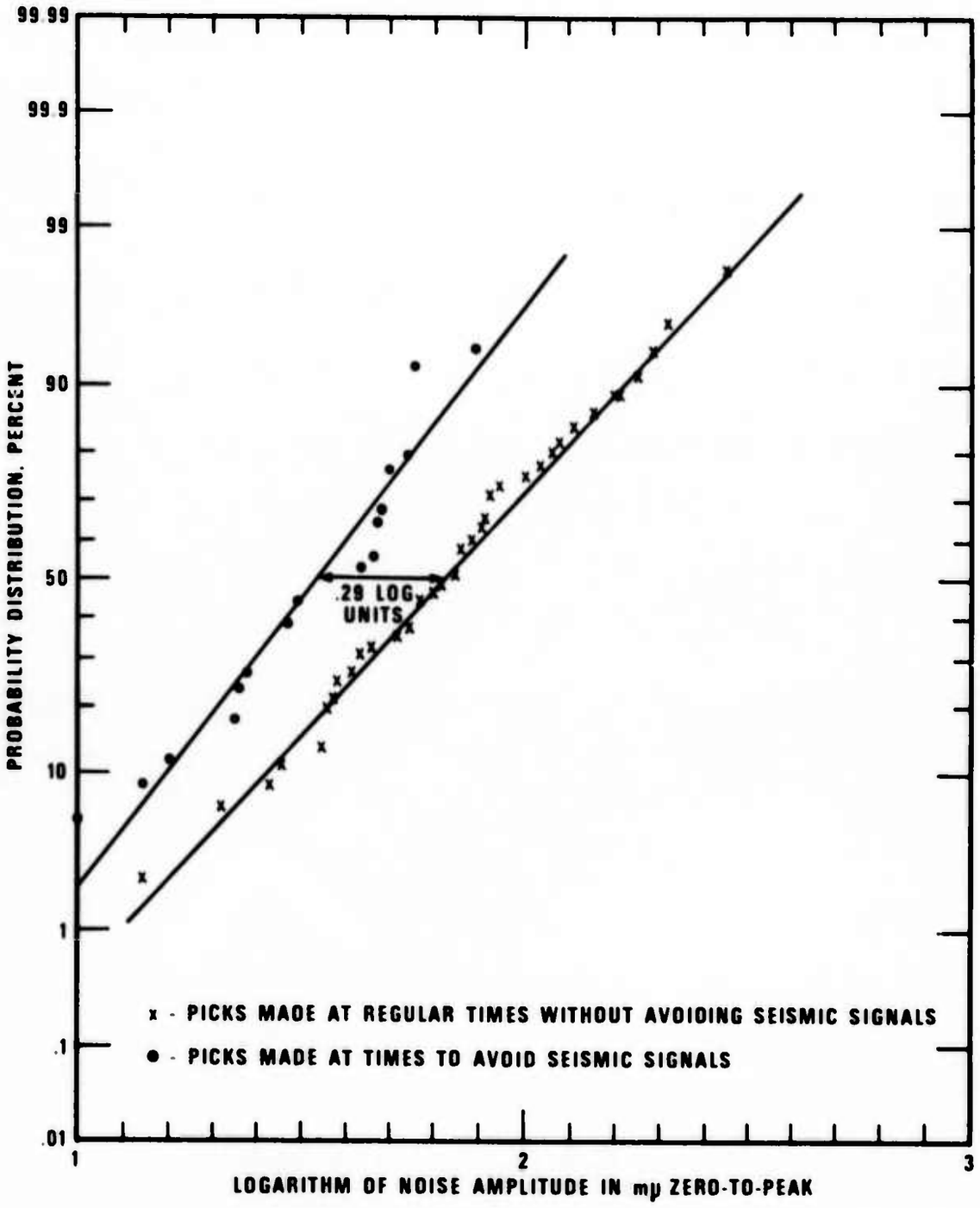


Figure 36. Distribution of CTA long-period noise amplitudes picked on the high-gain vertical component in 1972.

APPENDIX I

DERIVATION OF THE PROBABILITY DENSITY OF OBSERVED MAGNITUDE FOR A TWO-STATION NETWORK

APPENDIX I

Derivation of the Probability Density of Observed Magnitude for a Two-Station Network

When two stations detect events of fixed operational magnitudes m, the probability density of the combined observed magnitudes is given by the convolution of the individual densities if, as is clear from physical considerations, the individual densities are uncorrelated:

$$f_{\hat{m}_1+\hat{m}_2}(m_c) = \int_{-\infty}^{\infty} f_{\hat{m}_2}(m_c-\hat{m}_1) f_{\hat{m}_1}(\hat{m}_1)\, d\hat{m}_1 \qquad \text{(A1)}$$

where $m_c=\hat{m}_1+\hat{m}_2$. We will convert to average magnitude $\bar{\hat{m}}$ later.

The expression for $f(\hat{m})$ for fixed m at a single station has been given by Herrin and Tucker (1972) as:

$$f(\hat{m}) = \frac{\Phi\left(\frac{\hat{m} - \mu'}{\sigma_n}\right)}{\Phi\left(\frac{m - \mu'}{\sigma}\right)} \frac{1}{(2\pi)^{1/2}\sigma_s} \exp\left[-\frac{1}{2}\left(\frac{\hat{m} - m}{\sigma_s}\right)^2\right] \qquad \text{(A2)}$$

Substituting (A2) in (A1) and rearranging the exponents results in:

$$f_{\hat{m}_1+\hat{m}_2}(m_c) = \frac{1}{2\pi\sigma_s^2 p^2} \exp\left[-\frac{1}{2}\left(\frac{\frac{m_c}{2} - m}{\sigma_s/2^{1/2}}\right)^2\right] \int_{-\infty}^{\infty} \phi\left(\frac{m_c - \hat{m}_1 - \mu'}{\sigma_n}\right)$$

$$\phi\left(\frac{\hat{m}_1 - \mu'}{\sigma_n}\right) \exp\left[-\frac{1}{2}\left(\frac{\frac{m_c}{2} - \hat{m}_1}{\sigma_s/2^{1/2}}\right)^2\right] \, d\hat{m}_1$$

where $p = \phi\left(\frac{m-\mu'}{\sigma}\right)$. Making the change of variable:

$$z = \frac{\hat{m}_1 - \frac{m_c}{2}}{\sigma_s/2^{1/2}}$$

and rearranging the terms in the cumulative normal probability functions results in:

$$f_{\hat{m}_1+\hat{m}_2}(m_c) = \frac{1}{2\pi^{1/2}\sigma_s p^2} \exp\left[-\frac{1}{2}\left(\frac{\frac{m_c}{2} - m}{\sigma_s/2^{1/2}}\right)^2\right] \tag{A3}$$

$$\int_{-\infty}^{\infty} \phi\left(\frac{\frac{m_c}{2} - m}{\sigma_n} - \frac{\sigma_s z}{2^{1/2}\sigma_n}\right) \phi\left(\frac{\frac{m_c}{2} - m}{\sigma_n} + \frac{\sigma_s z}{2^{1/2}\sigma_n}\right) \frac{e^{-z^2/2}}{(2\pi)^{1/2}} \, dz$$

Using the substitutions

$$\gamma = \frac{\frac{m_c}{2} - m}{\sigma_n} \qquad \text{and} \qquad \tau = \frac{\sigma_s}{2^{1/2}\sigma_n} ,$$

we express the integral in (A3) as:

$$I = \int_{-\infty}^{\infty} \Phi\left(\frac{\gamma/\tau - z}{1/\tau}\right) \Phi\left(\frac{\gamma/\tau + z}{1/\tau}\right) \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz. \quad (A4)$$

Using the relation $1-\Phi(-x) = \Phi(x)$, we can separate the integral into two parts:

$$I = \int_{-\infty}^{\infty} \Phi\left(\frac{z + \gamma/\tau}{1/\tau}\right) \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz - \int_{-\infty}^{\infty} \Phi\left(\frac{z + \gamma/\tau}{1/\tau}\right) \Phi\left(\frac{z - \gamma/\tau}{1/\tau}\right) \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz. \quad (A5)$$

A theorem from Ellison (1964) provides the evaluation of the first integral in (A5):

$$\int_{-\infty}^{\infty} \Phi\left(\frac{z + \gamma/\tau}{1/\tau}\right) \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz = \Phi\left[\frac{\gamma}{(\tau^2 + 1)^{1/2}}\right] \quad (A6)$$

Another theorem from Zachs and Even (1966) provides the evaluation of an integral similar to the second one in (A5), except that their means in the cumulative normal distributions are of the same sign as well as being equal; an obvious substitution though of a positive for a negative sign can be carried through their derivation to produce:

$$\int_{-\infty}^{\infty} \Phi\left(\frac{z + \gamma/\tau}{1/\tau}\right) \Phi\left(\frac{z - \gamma/\tau}{1/\tau}\right) \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz =$$

$$\Phi\left[\frac{-\gamma}{(\tau^2 + 1)^{1/2}}, \frac{\gamma}{(\tau^2 + 1)^{1/2}}; \frac{1}{\tau^2 + 1}\right] \qquad \text{(A7)}$$

where $\Phi(-\mu, \mu, \rho)$ is the bivariate normal distribution with means $-\mu$ and μ and correlation coefficient ρ.

Expanding the bivariate normal distribution in terms of derivatives of the normal density function gives:

$$\Phi\left[\frac{-\gamma}{(\tau^2 + 1)^{1/2}}, \frac{\gamma}{(\tau^2 + 1)^{1/2}}; \frac{1}{\tau^2 + 1}\right] = \phi\left[\frac{-\gamma}{(\tau^2 + 1)^{1/2}}\right] \cdot \Phi\left[\frac{\gamma}{(\tau^2 + 1)^{1/2}}\right] + \sum_{n=0}^{\infty} \left\{ \frac{Z^{(n)}\left[\frac{-\gamma}{(\tau^2 + 1)^{1/2}}\right] Z^{(n)}\left[\frac{\gamma}{(\tau^2 + 1)^{1/2}}\right]}{(n + 1)!} \left(\frac{1}{\tau^2 + 1}\right)^{n+1} \right\}. \qquad \text{(A8)}$$

The product of even-ordered derivatives in the above series is positive while that of odd-ordered derivatives is negative, so that:

$$\Phi\left[\frac{-\gamma}{(\tau^2+1)^{1/2}}, \frac{\gamma}{(\tau^2+1)^{1/2}}; \frac{1}{\tau^2+1}\right] = \Phi\left[\frac{-\gamma}{(\tau^2+1)^{1/2}}\right]$$

$$\Phi\left[\frac{\gamma}{(\tau^2+1)^{1/2}}\right] + \sum_{n=0}^{\infty}\left\{\frac{(-1)^n\left\{Z^{(n)}\left[\frac{\gamma}{\tau^2+1)^{1/2}}\right]\right\}^2}{(n+1)!}\right.$$

$$\left.\left(\frac{1}{\tau^2+1}\right)^{n+1}\right\}. \qquad \text{(A9)}$$

Now using equations A9, A7, and A6 in A4, we can express the integral as

$$I = \left\{\Phi\left[\frac{\gamma}{(\tau^2+1)^{1/2}}\right]\right\}^2 - \sum_{n=0}^{\infty}\left\{\frac{(-1)^n\left\{Z^{(n)}\left[\frac{\gamma}{(\tau^2+1)^{1/2}}\right]\right\}}{(n+1)!}\right.$$

$$\left.\left(\frac{1}{\tau^2+1}\right)^{n+1}\right\}. \qquad \text{(A10)}$$

This expression goes asymptotically to 0 as $\gamma \to -\infty$ and asymptotically to 1 as $\gamma \to \infty$; its value at $\gamma=0$ and $\tau=1$ is $\frac{1}{6}$. The series summation is rapidly convergent for all γ and all τ. Expressing γ and τ in terms of the original variables and then substituting (A10) in (A3) gives the density function for m_c:

$$f_{\hat{m}_1+\hat{m}_2}(m_c) = \frac{1}{2(\pi)^{1/2}\sigma_s p^2} \exp\left[-\frac{1}{2}\left(\frac{\frac{m_c}{2} - m}{\sigma_s/(2)^{1/2}}\right)^2\right] \cdot$$

$$\left\{\left[\Phi\left(\frac{\frac{m_c}{2} - m}{\sigma'}\right)\right]^2 - \sum_{n=0}^{\infty} \frac{(-1)^n \left[Z(n)\left(\frac{\frac{m_c}{2} - m}{\sigma'}\right)\right]^2}{(n+1)!}\right.$$

$$\left.\left(\frac{\sigma_n^2}{\sigma'^2}\right)^{n+1}\right\}. \qquad \text{(A11)}$$

where $\sigma' = (\sigma_n^2 + \sigma_s^2/2)^{1/2}$. If we define $\bar{\hat{m}} = m_c/2$, then by probability axiom:

$$f(\bar{\hat{m}}) = 2 f_{\hat{m}_1+\hat{m}_2}(2\bar{\hat{m}}).$$

Applying this to (A11) gives the desired probability density function for observed average magnitude when both stations of a two-station network detect:

$$f(\hat{\bar{m}}) = \frac{\exp\left[-\frac{1}{2}\left(\frac{\hat{\bar{m}} - m}{\sigma_s/2}\right)^2\right]}{\pi^{1/2}\sigma_s\left[\Phi\left(\frac{m - \mu'}{\sigma}\right)\right]^2} \cdot \left\{\left[\Phi\left(\frac{\hat{\bar{m}} - m}{\sigma'}\right)\right]^2 - \sum_{n=0}^{\infty} \frac{(-1)^n\left[Z^{(n)}\left(\frac{\hat{\bar{m}} - m}{\sigma'}\right)\right]^2}{(n+1)!}\left(\frac{\sigma_n^2}{\sigma'^2}\right)^{n+1}\right\}. \quad \text{(A12)}$$

APPENDIX II

New NETWORTH Capabilities for Handling Source Bias and Noise Correlation

A. IDENTIFICATION

Title: NETWCORR

Programmer: Modification by H. Husted of NETWORTH by Mark Wirth

Date: 20 September 1973

B. PURPOSE

To compute the network magnitude bias for the threshold magnitudes computed by the standard NETWORTH program. Also to evaluate the effect on thresholds of noise and signal amplitudes which are correlated between stations.

C. USAGE

1. This is a FORTRAN IV Program for the 360/44. The principle program writeup is that of Wirth (1970) and having it at hand is necessary in order to use this one.

2. Modifications to NETWORTH control card formats:

If bias calculations are desired, a 1 is placed in column 80 of the A card. If in addition, the run is to be made with correlated signal and/or noise; a 2 is placed in column 80. Bias will be computed in this case also.

If there is a 1 or 2 in column 80 of the A card, then immediately following must come a card with (NSD, NRM, NRANDS, NREP, NS, NN, NREPK) (7I10). NSD is the number of stations which must detect before a magnitude is calculated. If correlated signal and/or noise are to be considered, this is reset internally to the number of stations required for detection, found in columns 9 and 10 of the A card. NRM is the number of randomly generated event threshold magnitudes which must be calculated before the average bias is computed. NRANDS is a random number to start the random number generator. Default to 77773777. These are the only additional parameters needed for the case of a simple bias calculation.

If one wishes also to evaluate the effect of correlated noise and/or signal; then the meaning of NRM changes slightly and the additional parameters NREP, NS, NN and NREPK, together with covariance or correlation matrices of the signal and noise must be read in.

NREP is the number of events which are randomly generated at each epicenter in an attempt to have NRM of them detected so that a reliable probability of detection estimate may be obtained. NREPK is the maximum number of times to repeat NREP trials. (A limit must be set since, for low magnitudes, detections may not occur even after large amounts of computer time have been used. Some care in choosing NREP and NREPK is needed to minimize the probability biases which

would result from cutting off the calculations just when NRM detections have been achieved.) NS and NN are 0, 1, or 2 as the signal and noise are: a) uncorrelated; b) correlated and covariance matrix is to be read in ; or c) correlated and correlation matrix is to be read in. The lower triangular portion of the covariance or correlation matrices are read in, signal matrix first. (See the attached data forms.) If NS or NN=0, the corresponding diagonal matrix must not be read in.

3. Timing

The program, especially with correlated noise and signal, can be very time-consuming. Restriction of magnitude range and careful selection of epicenters, together with preliminary test runs is recommended.

4. In general all options of NETWORTH and NETPLOT are retained with the exception that one cannot do subsets with correlated signal and/or noise. A default to the standard NETWORTH program occurs if all new parameters are zero.

D. METHOD

If a single bias calculation is desired a magnitude differential Δm_k^{ij} is generated for station i, epicenter j.

$$\Delta m_k^{ij} = N(0,\sigma_{is}^2)$$

where σ_{is} is the standard deviation of the signal at

station i. Then if A_{ij} is the amplitude for the threshold magnitude for station i, epicenter j and μ_i is the noise level we compute

$$\Phi(x_k^{ij}) \text{ where } x_k^{ij} = \left(\frac{\log_{10} A_{ij} + \Delta m_k^{ij} - (\mu_i + \log r)}{\sigma_{in}} \right)$$

where Φ is the cumulative normal probability distribution, σ_{in} is the noise standard deviation, and r is the signal-to-noise ratio required for detection.

Then a second random number U_k, uniform on the interval $0<U_k<1$ is generated, and if $U_k<\Phi^{ij}_k$ we detect and add Δm^{ij}_k to the sum over detecting station:

$$\sum \Delta m^{ij}_k.$$

After all stations have been considered, we divide by the number of stations detecting if $\geq$NSD yielding Δ_{mb}. Repeat until NSD or more stations have detected NRM times and save the average $(1/NRM)\sum \Delta m_b$ as the bias for epicenter j.

If noise and/or signal correlation exists then for each "random event" sets of random numbers with the required mean and covariance structure are generated by the methods outlined by Shumway and Blandford (1970). Then detection occurs at each station if $\{\log A+\Delta m-(\mu+\log r)\}$ is greater than zero. Sufficient iterations may be taken, as outlined in C-2 above to ensure stable estimates of probability of detection and bias.

E. REFERENCES

Shumway, R. H. and Blandford, R. R., 1970, A simulation of seismic discrimination analysis, Seismic Data Laboratory Report No. 261, Teledyne Geotech

von Seggern, D. H. and Blandford, R. R., 1974, Seismic threshold determination, TR 1974 - Teledyne Geotech

Wirth, M. H., 1970, Estimation of network detection and location capability, Research Memorandum, Teledyne Geotech.

F. DATA CARD FORMATS.

Attached.

DATA FORM NETWCORR	PAGE 1	NAME
		DATE

ID ALPHA	NOT USED
A	
L	NO CHANGE FROM NETWORTH
S	NO CHANGE FROM NETWORTH
U	NO CHANGE FROM NETWORTH

1 or 2

ESD A77 2/64

DATA FORM NETWCORR	PAGE 2	NAME
		DATE

THESE CARDS COME AFTER U CARD ONLY IF COLUMN 80 OF A CARD ≠ 0.

				0 uncorrelated	0 uncorrelated		
NSD - Stations Required for det. for Bias	NRM - No. Rep. to determine Avg Mag Bias and ~~probability~~ of det.	NRANDS Random Start	NREP - No. of trials/set (the min. no.)	NS Signal 1-covariance 2-correlation	NN Noise 1-covariance 2-correlation	NREPK - Max. no. of times to repeat simulation	NREP+NREPK not used unless col. 80 of Acard ≠ 2
Con(1,1) conn. on cov. of station 1			ONLY INPUT THE LOWER TRIANGULAR				
Con(2,1) conn. on cov of stat. 1-stat. 2	Con(2,3) conn. on cov stat. 2		PORTION OF THE MATRIX				
Con(3,I) stat. 3-stat. 1	Con(3,2) stat. 3-stat.2	Con(3,3) stat. 3					
Con(4,1) stat. 4-stat. 1	Con (4,2) stat. 4-stat. 2	Con(4,3) stat. 4-stat.3	Con(4,4) stat. 4				

Input Signal matrix first.

DATA FORM NETWCORR	PAGE 3	NAME
		DATE

NO CHANGE TO DSP, DLP, T, V or E CARDS

REPEAT THIS SEQUENCE OF ID CARD → E CARD UNTIL ALL CASES HAVE BEEN RUN.

AFTER LAST 'E' CARD

(Any ALPHA)
END NETWORTH

P MUST HAVE END CARD FOLLOWED BY 'P' CARD TO PROPERLY TERMINATE PROGRAM.

ESD A77 2/64

APPENDIX III

Writeup for the Program MSBNET

1. Program Name - MSBNET

2. Purpose - Analysis of detection statistics of seismic networks - see reference.

3. Procedure:

Every network simulation begins with an initial set of events, each specified by its "true" magnitude M_i. The range of body wave magnitudes under consideration increases from a minimum of M1 (Input) in increments of 0.1 magnitude unit, and the expected number of events in an interval $[m_1,m_2]$ of these "true" values is given by:

$$E\{N(M)\} = \int_{m_1}^{m_2} Ae^{\alpha(M1-M)}\,dm \qquad (1)$$

$A, \alpha > 0$

where A and α are input parameters. The total number of events, A/α, must not exceed 5000. The M_i are generated by seeing if a random number, uniform on the interval (0,1) is greater than the total number of expected events up to M_i divided by M_T, the total number expected, and less than the expected number up to M_{i+1} divided by M_T. If so, one event of magnitude M_i is generated. Once the total number of expected events has been generated, the program proceeds. Of

course some magnitude cells will be overfilled and some underfilled, just as with real seismicity data.

Once the sample seismicity has been determined, the following steps are carried out for each station:

1. Scatter (due to variations in source regions, transmission paths, etc.) is simulated by generating a randomized population of station amplitudes from the initial magnitude set as follows:

$$\hat{M}_{s_k} = am_i + b + \gamma_k(0, \sigma^2_{ss}) + \varepsilon_k(0, \sigma^2_{ms})$$

$$\hat{m}_{b_k} = m_i + \delta_k(0, \sigma^2_{sb}) + \eta_k(0, \sigma^2_{mb}) \qquad (2)$$

$$i = 1, 2, \ldots, 100$$

$$k = 1, 2, \ldots, N(M_i)$$

where a is the slope and b is the intercept of the true M_s-m_b relation and the γ_k, δ_k, ε_k, and η_k are random normal variates with the indicated means and variances. The standard deviations σ_{sb} and σ_{ss} may vary as input parameters from station to station, while σ_{ms} and σ_{mb} are constant. γ and δ are different for every event, but the same for every station. That is, ε and η are the "source bias" and γ and δ are the "path-station bias".

2. Least-squares regression lines on $\hat{m}_b$ and $\hat{M}_s$ are fitted to the values obtained from equation (2). In addition, a maximum likelihood linear fit (see writeup of program MAXLIK) is computed for $\hat{M}_s$ - $\hat{m}_b$ which allows for the presence of normally distributed errors in both variables. The slopes and intercepts are printed and also accumulated to provide averages over all stations.

3. The population is searched to find the number of events scattered into each $\hat{M}_s$ magnitude interval:

$$N_{s_i} : (aM_i - .05) \leq \hat{M}_s < (aM_i + .05) \qquad i = 1, 2, \ldots, 100$$

then again on $\hat{m}_b$:

$$N_{b_i} : (M_i - .05) \leq \hat{m}_b < (M_i + .05) \qquad i = 1, 2, \ldots, 100$$

These numbers are printed, stored individually, and also accumulated over all stations.

4. To provide a detailed measure of the shape of the scattered population, the average $\hat{m}_b$ is computed for each fixed value of $\hat{M}_s$. These averages are printed and accumulated. In addition, the average $\hat{M}_s$ for fixed $\hat{m}_b$ is computed, printed, and accumulated over all stations.

5. Station detection of events in the scattered population is simulated by first computing the probability of detection of each event based on $\hat{m}_b$:

$$P_{b_k} = \Phi\left[(\hat{m}_{b_k} - m_{b_o} - \log_{10}R)/\sigma_{n_B}\right] \qquad (3)$$

$$k = 1, 2, \ldots, N(m_{b_i})$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$

m_{bo} = m_b detection threshold (50% level) for the station under consideration.

R = signal to noise ratio required for detection (same for all stations).

σ_{n_B} = standard deviation of the noise rms for m_b for the station under consideration.

Then for each of the P_{b_k}, a random number ξ_k uniformly distributed between 0 and 1 is computed, and the two are compared. If $\xi_k > P_{b_k}$ the event is considered "not detected", and in network type I the $(\hat{m}_b, \hat{M}_s)$ pair is eliminated from the population. If, in network type I, the station fails to detect any events on $\hat{m}_b$, a message is printed and that station is omitted from the computation of network detection statistics.

In Network type I a station will not detect M_s unless it has detected m_b. This might correspond to a network made up from the bulletins of 10 LASA's. Network type II is more common in which M_s will be looked for at each station if there is a network detection on m_b. In network type II we retain the $(\hat{m}_b, \hat{M}_s)$ pair for further analysis even if there is no detection

on m_b. Also at this point, in preparation for network detection statistics, if the J^{th} event is detected on m_b the element NOMBD (J) is increased by 1 and the array AMBM (J) is increased by $\hat{m}_b$.

After this "weeding" process is completed, this new population of "observed" events is searched as in step 3 to determine the number of events remaining in each magnitude interval. This search is based once on the $\hat{M}_s$ values of the events, and again on the $\hat{m}_b$ values. The data are printed, stored individually, and also accumulated over all stations.

6. For each station the events observed by detection on $\hat{m}_b$ at that station (or all events in the case of network type II) are subjected to detection on $\hat{M}_s$ by the same method used in step 5. After this second "weeding" process, the remaining population is again searched for the number of events left in each magnitude interval. These data are printed, stored individually and accumulated over all stations. Again, in preparation for netwrok statistics the arrays NOMSD(J) and AMSM(J) are increased by 1 and $\hat{M}_s$ respectively if detection occurs on $\hat{M}_s$.

7. Least-squares regression lines are fitted to the doubly detected events, either to the entire remaining population, or to a truncated portion defined by:

$$M11 \leq \hat{m}_b \leq M22 \qquad MS11 \leq \hat{M}_s \leq MS22$$

where the limiting values are specified input parameters.

Regression of both $\hat{M}_s$ on $\hat{m}_b$ and $\hat{m}_b$ on $\hat{M}_s$ are performed; moreover, a maximum likelihood linear fit is computed for M_s-m_b as in step 2. The slopes and intercepts are printed and accumulated for network averages.

8. Step 4 is repeated for the observed population.

9. The ratio of the number of events detected on both $\hat{m}_b$ and $\hat{M}_s$ to the number detected on $\hat{m}_b$ alone is calculated and printed for each magnitude interval, first using $\hat{M}_s$ as the magnitude of each event, and again using $\hat{m}_b$. These ratios represent the probability of detection on $\hat{M}_s$ given prior detection on $\hat{m}_b$ (according to type I or type II network) as a function of $\hat{M}_s$ and $\hat{m}_b$ respectively at the given station.

10. Finally, a printer plot of $\hat{M}_s$-$\hat{m}_b$ for the events detected on both m_b and M_s is output for the station. This plot also shows the true M_s-m_b line.

Steps 1-10 are repeated for each station in the network. After all stations have been simulated, the accumulated data is presented in a network summary for the first set of events:

Station Summary

I. Over all initial "scattered" populations:

   1. Average slope and intercept of regressions on $\hat{m}_b$.
   2. Average slope and intercept of regressions on $\hat{M}_s$.

3. Average slope and intercept of maximum likelihood fits of $\hat{M}_s$-$\hat{m}_b$.

4. Accumulated number of events in each magnitude interval of $\hat{M}_s$, both incremental and cumulative.

5. Accumulated number of events in each magnitude interval of $\hat{m}_b$, both incremental and cumulative.

6. Average $\hat{m}_b$ for fixed $\hat{M}_s$.

7. Average $\hat{M}_s$ for fixed $\hat{m}_b$.

II. After detection on $\hat{m}_b$:

8. Accumulated number of events in each magnitude interval of $\hat{M}_s$ (same as I-4 in the case of network type II).

9. Accumulated number of events in each magnitude interval of $\hat{m}_b$.

III. After detection on both $\hat{m}_b$ and $\hat{M}_s$ (or on $\hat{M}_s$ only in the case of network type II):

10. Accumulated number of events in each magnitude interval of $\hat{M}_s$.

11. Accumulated number of events in each magnitude interval of $\hat{m}_b$.

12. Average $\hat{m}_b$ for fixed $\hat{M}_s$.

13. Average $\hat{M}_s$ for fixed $\hat{m}_b$.

14. The probability of detection on $\hat{M}_s$ given prior detection on $\hat{m}_b$ (or without consideration of detection on m_b for network type II) is obtained as a function of $\hat{M}_s$ by computing the ratio of the data of step 10 to the data of step 8 for each magnitude interval.

15. The probability of detection on $\hat{M}_s$ given prior detection on $\hat{m}_b$ (same as above) is obtained as a function of $\hat{m}_b$ by computing the ratio of the data of step 11 to the data of step 9 for each magnitude interval.

16. Average over stations of slope and intercept of regression fits of $\hat{M}_s$ on $\hat{m}_b$.

17. Average over stations of slope and intercept of regression fits of $\hat{m}_b$ on $\hat{M}_s$.

18. Average over stations of slope and intercept of maximum likelihood fits of $\hat{M}_s$ and $\hat{m}_b$.

Network Summary

19. The probability of network detection on $\hat{m}_b$ by NRMBD or more stations as a function of m_b, is calculated by seeing if NOMBD(J)$\geq$NRMBD (J^{th} event) and if so incrementing by 1 the NMBNTD cell determined by SQGB=AMBM(J)/NOMBD (J) = $\hat{m}_b$. If NOMBD(K) < NkMBD go to next event. After going through all events, NMBNTD(K)/N (m_k) is printed out.

20. For each event, after (19) if NOMBD(J)>NRMBD, if NOMSB≥NRMSD then we increment by 1 the NMSNTD (K) determined by SQGS=AMSM(K)/NOMSD (K) = $\hat{M}_s$ and the NMBMSD (K) cell determined by SQGB. After going through all events NMSNTD (K)/N (m_k) is printed out. Note that because of magnitude scatter the quantities printed out in (19) and (20) may be > 1.0.

21. To obtain the probability of detection of LR by at least NRMSD stations as a function of $\tilde{\hat{m}}_b$ we compute and print out NMBMSD (K)/ NMBNTD (K). These numbers are always ≤ 1.0.

22. NMBNTD (K), NMSNTD(K), and NMBMSD (K) are printed out, together with their cumulative versions.

23. Max-like and m_b and M_s regression lines are computed on the surviving ($\tilde{\hat{m}}_b$, $\bar{\hat{M}}_s$) network pairs for NSTAT (up to 10) sets of $\bar{\hat{M}}_s$ and $\tilde{\hat{m}}_b$ magnitude intervals (NSTAT) x (EM1, EM2; EMS1, EMS2).

24. The surviving $\tilde{\hat{m}}_b$, $\bar{\hat{M}}_s$ are printed out and plotted on the printer.

25. The steps 19-24 are repeated for NRDT (up to 10) pairs of (NRMBD,NRMSD) values.

Final Summary

26. The entire program loops NREP times (K=1, NREP); saving and accumulating the max-like slopes and intercepts as well as NMBNTD (K), NMSNTD (K), and NMBMSD (K) in KMBNTD (K, L) KMSNTD (K, L), and KMBMSD (K, L)(L=1, NRDT). When all repetitions are completed, average regression lines are printed, along with the ratio KMBMSD (K, L)/KMBNTD (K, L). KMBNTD (K, L), KMSNTD (K, L), and KMBMSD (K, L) are also printed, both incremental and cumulative. Finally, NRDT printer plots showing the accumulation of all previous network plots for each L are computed and printed.

Correlated Noise

27. If NCORR $\neq$ 0 a preliminary run of program SIMCRN is required to generate noise values with the required correlation matrix. These are output to a tape which must be mounted on SYS1 when run on the 360/44. The program is iterated as many times (NREP) as it is desired to iterate PROGRAM MSBNET. Input for program SIMCRN is described at the end of this writeup.

4. INPUT: Six general data cards followed by NSTAT cards for different data sets for regression calculations and by one data card for each station

(max. 30 stations) are required for each model run of program MSBNET. The variables and formats of the cards are as follows:

CARD 1 (6I10, 4I5)

IX — Initial value for random number generation. (If zero, IX is set to 777773777.)

IPNT — If zero, omit printout of individual station data. Network summaries and the final summary will still be printed.

IHPKL — If zero, omit printout of individual M_s, m_b event pairs for each station.

JIPKL — If zero, omit printout of M_s, m_b event pairs for network.

JZ3Z — If zero, network type I. If one, network type II.

NRDT — Number of (NRMBD,NRMSD) pairs to be used (see cards 3 and 4).

MBMIN, MBMAX, MSMIN, MSMAX — Limits on the m_b and M_s axes of the printer plots of station and network data. Default to 2.0-7.0 on both axes. If MBMIN = -1, the maximum and minimum data values will be used.

CARD 2 (6F10.2, 2I10)

M1 — Smallest "true" m_b value used in the generation of each population of "observed" events.

A, ALP — Parameters of the seismicity density function (see eq. 1) must be adjusted to give no more than 5000 events: A/ALP<5000.

BETA, CINT — Slope and intercept of the "true" M_s-m_b relation: M_s = BETA * m_b + CINT.

R — Signal to noise ratio required for detection.

N — Number of stations (must not exceed 30).

NREP — Number of repetitions.

CARD 3 (10I5)

NRMBD(I) — Minimum number of station detections of $\hat{m}_b$ required to declare a network detection. (I = 1, NRDT; see card 1)

CARD 4 (10I5)

NRMSD(I) — Minimum number of station detections of $\hat{M}_s$ required to declare a network detection. (I = 1, NRDT; see card 1)

CARD 5 (4F10.4, I10)

SGNTB — Standard deviation of network m_b for maximum likelihood regressions.

SGNTS — Standard deviation of network M_s for maximum likelihood regressions.

SIGSS — Standard deviation for m_b source bias.

SIGBS — Standard deviation for M_s source bias.

NCORR — Switch ($\neq$ 0) for a correlated noise and signal run.

CARD 6 (4F10.2, I10)

M11 Lowest magnitude m_b to consider for regression fit of detected events.

M22 Highest magnitude m_b to consider for regression fit of detected events.

MS11 Lowest magnitude M_s to consider for regression fit of detected events.

MS22 Highest magnitude M_s to consider for regression fit of detected events.

NSTAT Number of sets of EM1, EM2, EMS1, EMS2 to follow. If NSTAT = 0, station data cards come next.

CARD 7 - CARD (6 + NSTAT) (4F10.2)

EM1 (I)
EM2 (I)
EMS1(I)
EMS2(I)
I = 1, NSTAT

Truncation limits on m_b and M_s for additional regression fit (same definitions respectively as Card 6).

I^{th} STATION DATA CARD (I = 1, N) (5F10.4)

AMBO(I) m_b detection threshold for the I^{th} station, 50%.

AMSO(I) M_s detection threshold for the I^{th} station, 50%.

ASGMB(I) Standard deviation of observed events about the M_s-m_b relation along the m_b axis for the I^{th} station, with variance due to source bias excluded.

ASGMS(I) Standard deviation of observed events about the M_s-m_b relation along the M_s axis for the I^{th} station with variance due to source bias excluded.

ASGB(I) Standard deviation of the noise rms for m_b at the I^{th} station.

ASGS(I) Standard deviation of the noise rms for M_s at the I^{th} station.

5. Error Messages:

Self-explanatory.

6. Memory requirements:

200 K on 360/44. Makes extensive use of the auxiliary disc memory.

7. References: Seismic Threshold Determination, TR-74-3, D. H. von Seggern, R. R. Blandford, 1974.

8. Programmers: H. Husted, T. McElfresh, R. Blandford.

```
C     PROGRAM SIM CRN CORRELATED DATA FOR NETWORK
C   PROGRAM TO GENERATE 5000 CORRELATED OBSERVATIONS FOR 10 STATIONS
      DIMENSION CORR(10,5000),CCVAR(10,10),AMEAN(10),X(10)
      DIMENSION Y(10)
C
C
C   1-5 IFCOR= 1 COVARIANCE MATRIX IS INPUT
C            = 2 CORRELATION MATRIX IS INPUT
C
C   6-10  N     NUMBER OF STATIONS USED IN SIMULATION (N)
C
C   11-20 NRAN  STARTING RANDOM NUMBER (DEFAULT TO 77773777)
C
C   21-30 NREP  NO. OF TIMES TO REPEAT SIMULATION
C
C
C
C
C   1-10  AMEAN(1)  THE MEAN OF THE FIRST STATION
C     .                      .                .
C     .                      .                .
C     .                      .                .
C   71-80  AMEAN(8)  THE MEAN OF THE 8TH STATION
C
C               REPEAT UNTIL ALL STATIONS MEANS ARE READ
C
C
C
C
C   1-5    STANDARD DEVIATION OF STATION 1
C     .                      .                .
C     .                      .                .
C     .                      .                .
C   76-80  STANDARD DEVIATION OF STATION 16
C
C            REPEAT UNTIL ALL STATION STANDARD DEVIATIONS ARE READ
C
C
C
C          ONLY INPUT THE LOWER TRIANGULAR PORTION OF MATRIX
C          CARD 1   AS(1,1)
C          CARD 2   AS(2,1),AS(2,2)
C          CARD 3   AS(3,1),AS(3,2),AS(3,3)
C            .        .                .
C            .        .                .
C            .        .                .
C     CARD NST AS(NST,1),AS(NST,2),AS(NST,3),. . .,AS(NST,NST)
C
C
```

March 1, 1973

A SUBROUTINE FOR MAXIMUM LIKELIHOOD LINEAR FITTING ON THE IBM 360-44

1. Subroutine Name: MAXLIK

2. Purpose:

This subroutine performs maximum likelihood fitting of a straight line to measured points having correlated and normally distributed errors in both variables. The method involves point adjustments along likelihood ellipse diameters conjugate to the fitted line.

3. Calling Sequence:

CALL MAXLIK (X, Y, N, G, R, A, B, SA, SB, SX, SY)

4. Input:

X = 1st variable array
Y = 2nd variable array
N = Number of data points in X, Y arrays
G = Ratio of standard deviations σ_y/σ_x
R = Correlation coefficient of errors in X and Y

5. Output:

A = Slope of the fitted line: Y = A* X + B.
B = Intercept of the fitted line: Y = A* X + B.
SA = Standard deviation of the slope for the uncorrelated (R = 0.) case. If $R \neq 0$, the returned value of SA is -1.0.

SB = Standard deviation of the intercept for the uncorrelated (R = 0.) case. If R ≠ 0, the returned value of SB is -1.0.

SX = Standard deviation of the X observations with respect to their true values.

SY = Standard deviation of the Y observations with respect to their true values.

6. Error Messages: None.

7. Memory Requirements: Less than 2K bytes

8. References:

Ericsson, U., Maximum likelihood linear fitting when both variables have normal and correlated errors; The Research Institute of National Defense, Stockholm 80, Sweden, FOA 4 Rapport, C4474-A1, Sept. '71.

```
      SUBROUTINE MAXLIK(X,Y,N,G,RHC,A,B,SA,SB,SX,SY)
C
      DIMENSION X(1),Y(1)
C
      TN=FLOAT(N)
C
      XB=0.0
      YB=0.0
      DO 30 J=1,N
      XB=XB+X(J)
      YB=YB+Y(J)
   30 CONTINUE
      XB=XB/TN
      YB=YB/TN
C
      SUU=0.0
      SUV=0.0
      SVV=0.0
      DO 40 J=1,N
      UJ=X(J)-XB
      VJ=Y(J)-YB
      SUU=SUU+(UJ*UJ)
      SUV=SUV+(UJ*VJ)
      SVV=SVV+(VJ*VJ)
   40 CONTINUE
C
      BX=SUV/SUU
      AX=XB-(BX*XB)
      BY=SUV/SVV
      AY=XB-(BY*YB)
      FAC=SQRT(1./(TN-2.))*((1./(BX*BY))-1.))
      SBX=BX*FAC
      SBY=BY*FAC
      G2=G*G
      SO=(RHO*BY)-G2
      S1=(G2/BX)-1./BY)
      S2=1.-RHO*G*(1./BX)
C
      A=(-S1+SQRT(S1*S1-4.*SO*S2))/(2.*S2)
      B=YB-A*XB
C
      SS=0.0
      DO 50 J=1,N
      F=Y(J)-B-A*X(J)
      SS=SS+(F*F)
   50 CONTINUE
      SX=SQRT(SS*(1./(G2+A*A-2.*A*RHO*G))*(1./(TN-2.)))
      SY=G*SX
C
      IF(RHO.EQ.0.0) GO TO 60
      SA=-1.0
      SB=-1.0
      RETURN
   60 DA=(G2+(BX/BY))/(TN-1.)
      DB=G2+A*A
      DC=(1./(BX*BY))-1.0
      SA=A*SQRT(DA*DB*DC)
      SB=SQRT((SVV/(TN*(TN-2.)))+(XB*XB*SA*SA))
C
      RETURN
      END
```